Distribution Agreement

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

Signature:

Phebe B. Kemmer	Date

Statistical Approaches for Exploring Brain Connectivity with Multimodal Neuroimaging Data

By

Phebe Brenne Kemmer

Doctor of Philosophy

Biostatistics

Ying Guo, Ph.D.
Advisor

F. DuBois Bowman, Ph.D.
Advisor

Jian Kang, Ph.D.
Committee Member

Helen Mayberg, Ph.D.
Committee Member

Accepted:

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

Date

Statistical Approaches for Exploring Brain Connectivity with Multimodal Neuroimaging Data

By

Phebe Brenne Kemmer
M.S., Emory University, 2013
B.S., University of Michigan, 2008

Advisors: Ying Guo, Ph.D. and F. DuBois Bowman, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Biostatistics
2016

Abstract

Statistical Approaches for Exploring Brain Connectivity with Multimodal Neuroimaging Data

By
Phebe B. Kemmer

Advances in neuroimaging technology provide a gateway for studying the function and structure of the human brain, which stands to improve our understanding of neural networks and yield important insights about brain disorders. The statistical analysis of neuroimaging data poses a challenge, because the data is high-dimensional and contains spatial and temporal correlations. The focus of this dissertation is to develop statistical methods for multimodal neuroimaging data that allow us to explore the relationship between functional and structural brain connectivity, and investigate how this relationship differs between healthy and diseased brains.

A set of functionally connected brain networks can be estimated from fMRI data using independent component analysis (ICA). However, this approach ignores information about the underlying structural connections, which are believed to facilitate functional connectivity between remote brain regions. For the first topic, we propose a novel measure of the strength of structural connectivity (sSC) underlying these functional networks, by incorporating structural information from DTI data. To conduct inference on our sSC measure, we estimate a covariance term that considers spatial similarity between observations via a parametric semivariogram model with a novel distance metric. We demonstrate the performance of our proposed measure with simulation studies, and apply our method to an fMRI and DTI dataset. We find that sSC is associated with component reliability, demonstrating the benefit of leveraging information from structural data in the estimation of functional networks from fMRI data.

The second and third topics propose statistical frameworks for modeling the relationship between functional and structural connectivity across the whole-brain network. The second topic presents a hierarchical model with a linear link function to describe the association at each edge in the network, and uses the EM algorithm to estimate the model parameters. We consider both correlation and partial correlation as measures of functional connectivity. The third topic considers a more flexible approach to modeling the function-structure association by using copulas. In this way, we can model the marginal distributions of functional and structural connectivity data, and separately, estimate their association using a copula function. For each method, we conduct simulation studies to evaluate performance, and apply the proposed methods to an fMRI and DTI dataset, demonstrating biologically meaningful findings.

Statistical Approaches for Exploring Brain Connectivity with Multimodal Neuroimaging Data

By

Phebe B. Kemmer
M.S., Emory University, 2013
B.S., University of Michigan, 2008

Advisors: Ying Guo, Ph.D. and F. DuBois Bowman, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Biostatistics
2016

Acknowledgments

First and foremost, I want to thank my advisors, Dr. Ying Guo and Dr. DuBois Bowman, for providing invaluable guidance and mentorship during my time at Emory. Their knowledge and insights have greatly enriched my Ph.D. experience and education.

I would also like to thank my dissertation committee members, Dr. Jian Kang and Dr. Helen Mayberg, for their thoughtful suggestions and feedback, which have significantly improved my dissertation research and presentation.

I am very grateful to be a member of Emory's Center for Biomedical Imaging Statistics (CBIS) research group, led by Dr. Ying Guo. This small community of students, postdocs, and faculty has helped foster my passion for neuroimaging research and provided a useful sounding board for discussing ideas. Thank you to CBIS alumni Anthony Pileggi, Dr. Shuo Chen, and Dr. Gordana Derado for patiently teaching me the intricacies of the cluster, shell scripting, and Matlab coding when I was first getting started with my dissertation work. I would also like to acknowledge CBIS members Yikai Wang, Dr. Subhadip Pal, and Dr. Ran Shi for lending their knowledge and assistance as I worked through some of the trickier parts of my dissertation methodology, and Dr. KiSueng Choi of the Mayberg lab for providing expert guidance on the proper handling and preprocessing of neuroimaging data.

Thank you to my Emory colleagues and friends, as well as to the BIOS department faculty and staff, for making my PhD experience so rewarding and enjoyable. I am very fortunate to have started my PhD journey with the cohort of Christina Mehta, Emily Mitchell, Pallavi Mishra-Kalyani, Shuling Liu, and Xiaoyan Sun. These women have been a tremendous source of support, inspiration, and friend-

ship throughout my time at Emory.

Lastly, none of this work would have been possible without the encouragement of my family and friends. Thank you to my parents, Connie Raab and Gordon Brenne, my in-laws Teri and Jeff Kemmer, and my siblings, Lara and Nik, for their unending love and enthusiasm. My most heartful gratitude goes to my best friend and husband, Zachary Kemmer, to whom this work is dedicated. I could not have completed this process without him.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

Advances in in-vivo neuroimaging technology have provided a gateway for researchers to study the structure and function of the human brain, which can yield important insights on the neurophysiology underlying both healthy and diseased brains. Various neuroimaging techniques have been developed to investigate different aspects of brain processes. For example, fMRI and PET imaging can be used to measure brain activations and functional connectivity, DTI considers the structural connections within the brain, and MRI is useful for distinguishing tissue types and detecting abnormalities. Multimodal neuroimaging methods aim to leverage these complementary measurements by combining two or more datasets acquired with different techniques. By integrating information across modalities, we might improve our understanding of brain structure and function.

The statistical analysis of neuroimaging data is challenging because it is high-dimensional and contains complex spatial and temporal correlations. The main objective of this dissertation work is to develop novel statistical methods that combine information across functional and structural neuroimaging modalities, in order to provide a comprehensive examination of the neural network and investigate the relationship between brain structure and function.

This dissertation is organized as follows: the remainder of Chapter 1 provides background information on the human brain, a description of the functional and structural neuroimaging data modalities under consideration (i.e. fMRI and DTI), a review of existing methods for neuroimaging data analysis, and an overview of the motivating dataset that we will use in our analyses. Chapter 2 presents a novel statistic for measuring the strength of structural connectivity underlying the functional networks estimated by independent component analysis (ICA). Chapter 3 discusses a modeling framework for assessing the relationship between functional and structural connectivity across the whole brain network, using a linear link

function. This work is extended in Chapter 4, which proposes a more flexible approach to modeling the function-structure association by using copulas.

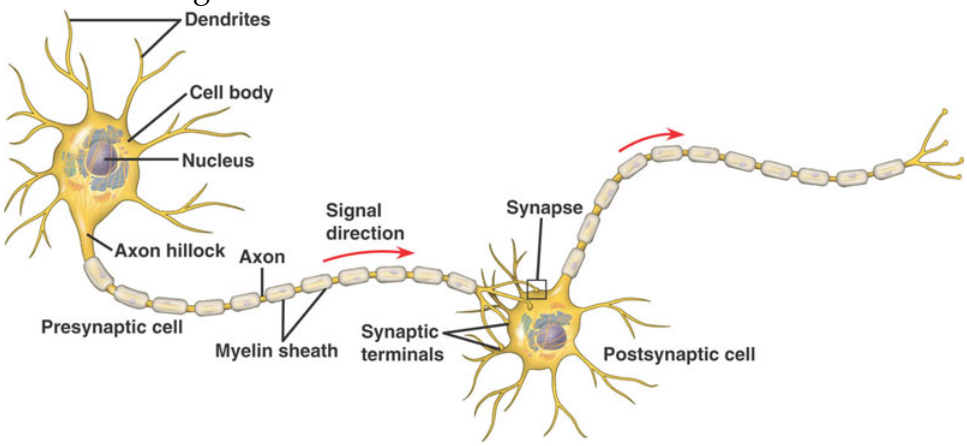
1.2 Organization of the Human Brain

The human brain is a complex organ, responsible for producing our thoughts and emotions, interpreting sensations, coordinating body movements, and much more. It weighs roughly 3 pounds and contains 100 billion nerve cells, known as neurons (Herculano-Houzel, 2009), Each neuron receives and sends signals to other neurons, enabling communication between remote brain locations. On average, each neuron is connected to 7,000 other neurons, yielding over 100 trillion total neural connections in the brain (Drachman, 2005).

The basic components of a neuron are: 1) the cell body, which is the control center of the neuron, 2) dendrites, which receive incoming messages from other nerve cells, and 3) axons, which send outgoing signals to other cells. Axons are sheathed in myelin, a fatty substance that facilitates the transmission of the neural signal. Neurons communicate with each other through electrical impulses called "action potentials" that fire when a nerve is stimulated. When a neuron fires, the impulse moves down the axon and transmits the signal to the postsynaptic neuron by releasing neurotransmitters in the synapse between the cells (see Figure 1.1). Generating this electrical/chemical signal transmission requires energy, provided by glucose and oxygenated-hemoglobin. Thus increased brain activity is associated with a higher rate of metabolism and increased oxygenated-hemoglobin concentration (Bear et al., 2007).

The brain consists of two hemispheres (the left and right), each of which has an outer layer of gray matter called the cerebral cortex, which is supported by an inner layer of white matter. The gray matter is composed of nerve cell bodies and den-

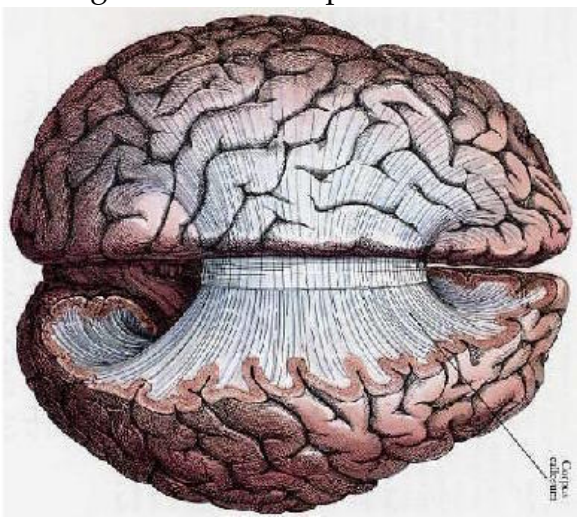
Figure 1.1: Communication between neurons



(source: http://vaxtruth.org/2011/10/offit-synapse/neuron-synapse/)

drites, while the white matter contains bundles of myelinated axons, called fiber tracts, that provide a physical link for transmitting signals between nerve cells in the cortex. The three main types of white matter tracts are: 1) projection fibers, which run vertically to connect the cortex to the rest of the body, 2) commisural tracts, which connect brain regions across the hemispheres (e.g. the corpus callosum, see Figure 1.2), and 3) association tracts, which connect brain regions within hemispheres (Bear et al., 2007).

Figure 1.2: The corpus callosum



(source: http://hubel.med.harvard.edu/book/b34.htm)

The hemispheres of the brain can be subdivided into four lobes: frontal, parietal, occipital, and temporal. Although they communicate with one another, each lobe is generally associated with different functions, e.g. the occipital lobe with vision, and the temporal lobe with auditory processes (Bear et al., 2007). To study the brain at a region level, however, we need to use a finer parcellation of the brain. Two common parcellation schemes are the Brodmann atlas, which defines 48 brain regions based on cytoarchitectural properties, and the Automated Anatomical Labeling (AAL) atlas, which defines 90 cortical brain regions based on the anatomical location of major sulci and gyri (the fissures and ridges of the cortical surface) (Tzourio-Mazoyer et al., 2002).

Although normal healthy brains have the same general organization, all brains exhibit slight differences from individual to individual. Individual subject's brains can differ in size by up to 30%. Therefore, for group neuroimaging studies, the brain images collected from different subjects must be transformed into a common coordinate space to accommodate individual differences in brain size and orientation. The Talairach and Montreal Neurological Institute (MNI) space are the most commonly used standard atlases for this spatial registration step. Talairach coordinates (Talairach and Tournoux, 1988) are based on the single brain of a 60-year old French woman, whereas the more recently developed MNI atlas is based on a large sample of structural images from healthy subjects (Evans et al., 1993), and is thus a more representative standard.

1.3 Functional Neuroimaging

Functional neuroimaging plays a large role in the study of neural networks; it enables researchers to localize brain activity in response to an experimental task and identify functionally connected brain areas. Functional Magnetic Resonance Imag-

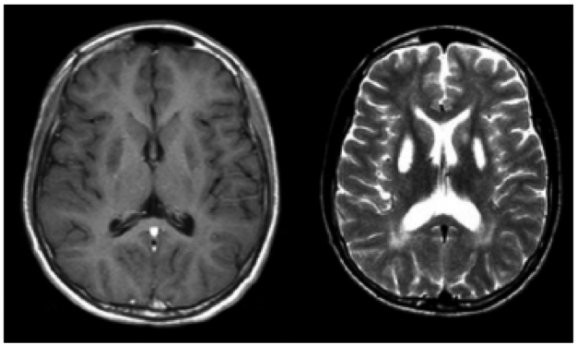
ing (fMRI), Positron Emission Tomography (PET), electroencephalography (EEG), and magnetoencephalography (MEG) are all functional neuroimaging techniques that measure correlates of brain activity in-vivo over time. fMRI and PET can detect the localized changes in cerebral blood flow associated with neural activity with relatively high spatial resolution, but the temporal resolution is limited by the slower rate of brain blood flow and blood oxygenation. EEG and MEG on other hand, detect the underlying electrical activity of the brain, which allows a higher temporal resolution than that of fMRI and PET, but results in poorer spatial resolution (Bear et al., 2007). For my dissertation work I focus on the functional modality of fMRI data, which is widely used in neuroimaging studies, and give an overview of this technique in the next section.

1.3.1 Basic Principles of Magnetic Resonance Imaging (MRI)

Since fMRI is an extension of MRI, we will first describe how MRI scans are acquired. MRI is a non-invasive imaging technique based on the principles of nuclear magnetic resonance (NMR), that captures images of hydrogen atom nuclei inside the brain. MRI relies on the fact that the human body is comprised primarily of water and fat, and thus hydrogen protons, which have nuclear spin. When a subject is placed in a large magnetic field (B_0), the hydrogen protons in the body tend to align parallel to the direction of B_0, producing a net magnetization in that direction. If a radio frequency (RF) pulse is applied at the resonance frequency, then the protons absorb the energy and jump to a higher-energy state (anti-parallel to B_0). After the RF pulse, some of the high-energy protons move back to their low-energy state (parallel to B_0), producing the MR signal measured by the coils. T1 relaxation defines the rate at which longitudinal magnetization recovers after the RF pulse. When the RF pulse is applied, all the protons are spinning in phase; after the RF pulse, the spins get out of phase at different rates, resulting in MR signal decay.

T2 and T2* relaxation measures this rate of decay of the transverse magnetization. Protons in different tissue types return to equilibrium at different T1 and T2/T2* relaxation rates, allowing us to visualize the contrast between gray matter, white matter, and CSF (Edelman and Warach, 1993). Figure 1.3 shows an example of the tissue contrast for T1- vs T2-weighted MRI.

Figure 1.3: T1- vs. T2-weighted MRI images



(Almeida et al., 2012)

The T1-weighted MRI provides an image with higher spatial resolution, while the T2-weighted image can be captured more rapidly. Thus, T1-weighting yields high-resolution structural MRI scans to be used for visualizing brain tissue, while T2-weighting is useful for capturing scans over time, i.e. functional MRI scans. An additional MRI modality, diffusion-weighted MRI, is useful for measuring structural connections in the brain, and will be discussed in section 1.4.1.

1.3.2 Functional MRI (fMRI) data

1.3.2.1 The fMRI BOLD signal

fMRI measures brain activity as a function of the hemodynamic (i.e. blood flow) response to experimental stimuli. As mentioned previously, firing neurons consume lots of energy, and this change in metabolism increases the need for oxygen, carried by hemoglobin. The human body will provide more hemoglobin than consumed

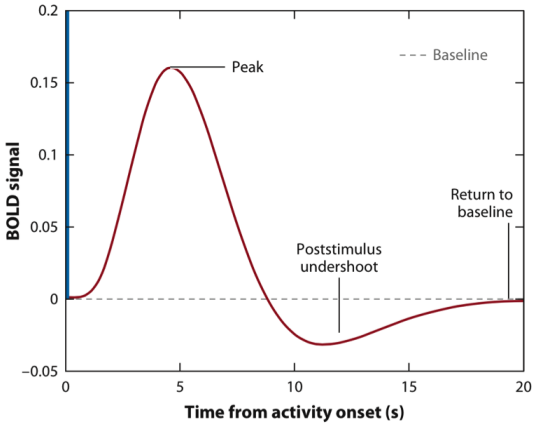
energy to the active brain area, thus increasing the concentration of oxygenated hemoglobin relative to that of de-oxygenated hemoglobin. Because de-oxygenated hemoglobin is more paramagnetic, it distorts the surrounding magnetic field, resulting in a weaker T2* signal where oxygen concentration is low, and a stronger T2* signal in the areas where oxygen concentration is high. fMRI uses this blood-oxygen-level-dependent (BOLD) contrast to measure differences in brain activity over time (Uludag et al., 2005).

The change in the fMRI BOLD signal that is triggered by an increase in neural activity is referred to as the hemodynamic response function (HRF; displayed in Figure 1.4). The HRF is characterized by a delayed response, since it onsets about 2 seconds after the neural activity, and peaks 5-8 seconds after neural activity has peaked. Another challenging property of the fMRI signal is that the magnitude of signal change in response to neural activity is quite small, roughly 0.5-3% for a 1.5T scanner. The fMRI signal contains many sources of noise, including noise from random neural activity, noise from the scanner due to drift and non-uniformities in the magnetic field, as well as physiological noise due to the subject's heartbeat, breathing, and head movement during the scan. These artifacts can be minimized by several preprocessing steps (see section 1.3.3), which should be applied to the data before conducting further analysis .

1.3.2.2 fMRI data structure

In an fMRI study, each subject is scanned multiple times over the course of an experimental session, so fMRI data can be represented as a series of 3D brain scans collected over time. Each of these 3D scans is comprised of small cubic units known as voxels (volumetric pixels), arranged in a large 3D matrix. The BOLD signal is measured at each voxel, resulting in an intensity value that represents the level of brain activity at that location. Plotting the BOLD signal from a particular

Figure 1.4: The Hemodynamic Response Function (HRF)



(Bowman, 2014)

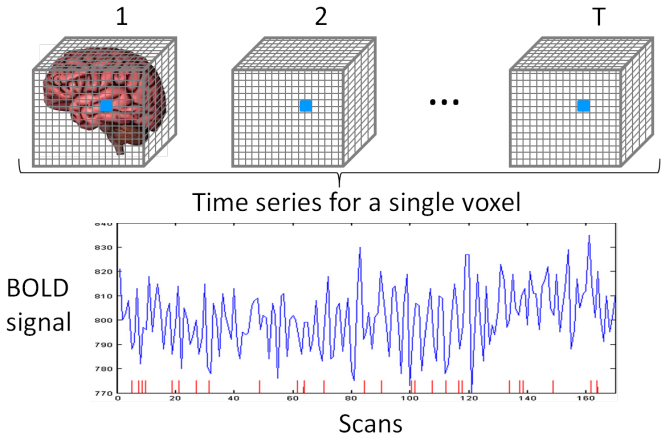
voxel at each scan time results in a single-voxel time course of brain activity (see Figure 1.5). For simplification, fMRI data can be rearranged into a VxT matrix, where V is the number of voxels, and T is the number of scans.

For an individual subject, an fMRI scanning session typically includes hundreds of thousands of voxels per 3D scan and hundreds of scans per session (usually recorded every 2-3 seconds). An entire fMRI study may consist of multiple subjects and repeated scanning sessions. The massive dimensionality of this data can pose computational burdens. Other complexities of fMRI data include spatial correlations between neighboring voxels, and temporal correlations between scans, which present challenges for statistical analysis.

1.3.2.3 Resting State

fMRI scans are acquired while a subject lies in the scanner, either performing some experimental task or at rest. In the resting state paradigm, subjects are often instructed to relax and to not think of anything in particular, or to stare at a fixation cross, while the spontaneous fluctuations in their brain activity are measured (Raichle, 2011). Rather than observing a brain that is silent at rest, studies

Figure 1.5: fMRI data structure



of healthy controls have consistently identified a set of "resting state networks" (Smith et al., 2009), most notably the Default Mode Network (Buckner et al., 2008). Since Biswal et al. (1995) first noted the presence of the motor network during rest, resting-state has become an increasingly popular method for studying the brain. Resting state can also be useful for exploring differences in functional networks between healthy and diseased brains (Greicius, 2008; Zhang and Raichle, 2010), and has the added benefit that it is easier to acquire data and compare results between studies.

1.3.3 fMRI Preprocessing Pipeline

Several preprocessing steps are necessary to prepare the fMRI data for further analysis.

1. **Brain extraction** to remove skull and non-brain matter.

2. **Slice timing correction** to account for timing differences in 2D slice acquisition.

3. **Motion correction** to re-align scans to a reference image.

4. **Co-registration** to align a subject's functional T2*-weighted scan to the high-resolution T1-weighted structural image.

5. **Spatial registration** maps each subject's brain to a standard coordinate space (i.e. the Talairach or MNI atlases).

6. **Spatial smoothing** increases the signal-to-noise ratio and blurs any residual anatomical differences that remain after spatial registration.

7. **Temporal smoothing** to remove low-frequency noise due to scanner drift, the subject's heartbeat and respiration, etc.

Preprocessing steps and the order in which they are performed are important because they can affect the spatial and temporal correlation structure of the data (Smith et al., 2004). Software tools that are commonly used to preprocess functional brain images include the Statistical Parametric Mapping (SPM) toolbox in Matlab, FMRIB Software Library (FSL), and Analysis of Functional NeuroImages (AFNI).

1.3.4 Functional Connectivity (FC) Analysis

The objectives of functional neuroimaging research are primarily studies of activation and functional connectivity. Activation studies aim to localize brain activity in response to experimental stimuli, and compare between subject subgroups (e.g. patients vs. controls, before vs. after treatment, etc.). The goal of functional connectivity studies, on the other hand, is to identify brain regions with similar patterns of brain activity.

Functional connectivity (FC) is defined as the temporal coherence between the BOLD signal of spatially remote brain regions (Richiardi et al., 2011; Friston, 1994). A set of brain regions with similar patterns of brain activity comprise a functionally connected brain network, called a **functional network**. Many brain disorders

are associated with disruptions in functional connectivity (e.g. Multiple Sclerosis, Alzheimer's Disease) (Zhang and Raichle, 2010).

Common methods for identifying functional networks include the seed-based correlation approach and data-driven partitioning algorithms, such as clustering, Principal Component Analysis (PCA), and Independent Component Analysis (ICA).

The seed-based approach selects a seed voxel from a region of interest (ROI), and computes the correlation between the seed voxel's time course and the time course of every other voxel in the brain, generating a connectivity map. An advantage of this approach is that it is simple to implement and the results have a straightforward interpretation. However, the choice of the seed voxel requires a priori knowledge of ROIs and can be subjective. It also ignores the network relationship between all voxels. Multivariate partitioning approaches, on the other hand, allow for an FC analysis of the whole brain simultaneously.

Clustering is a data-driven approach that groups voxels into clusters based on the similarity of their time courses, such that each cluster has a distinct neural pattern. Several different clustering algorithms, including K-means, fuzzy clustering, and hierarchical clustering, have been developed (Bowman et al., 2007).

Principal Component Analysis (PCA) and Independent Component Analysis (ICA) decompose the observed fMRI signal into components that represent functional networks. In PCA, the goal is to extract component maps that are uncorrelated and to summarize the variability in the observed data into as few components as possible (McKeown et al., 2003). ICA is an extension of PCA that decomposes the signal into statistically independent components. The stricter criteria for spatial independence used by ICA improves the accuracy of the estimated functional networks (McKeown et al., 2003). ICA has also been shown to be more effective than PCA at identifying functional networks in resting state (Beckmann and Smith, 2005), and runs much faster than clustering techniques (Meyer-Baese et al.,

2004). For these reasons, we will focus on the method of ICA for the FC analysis of resting-state fMRI data.

1.3.4.1 Independent Component Analysis (ICA)

Independent Component Analysis (ICA) was first described by Comon (1994) as a blind source separation method. A motivating example for ICA is the "cocktail party problem." Imagine a scenario in which two people talking simultaneously in a closed room, and there are two microphones placed in the room, each recording signals over time, $x_1(t)$ and $x_2(t)$. Each recorded signal is a weighted sum of the speech signals from the two speakers, $s_1(t)$ and $s_2(t)$. This can be expressed as:

$$x_1(t) = a_{11}s_1 + a_{12}s_2$$

$$x_2(t) = a_{21}s_1 + a_{22}s_2$$

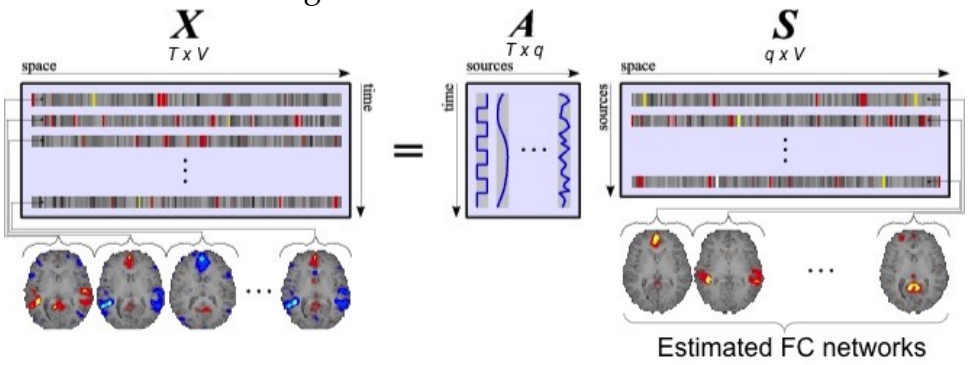
The goal of blind source separation methods like ICA is to estimate the original source signals $s_1(t)$ and $s_2(t)$, using only the recorded signals $x_1(t)$ and $x_2(t)$.

When applied to fMRI data, the goal of ICA is to decompose the observed fMRI signal into a set of q statistically independent component maps and their associated time courses (see Figure 1.6).

In Figure 1.6, $\boldsymbol{X}_{TxV}$ is the matrix of observed fMRI data; $\boldsymbol{A}_{Txq}$ is the "mixing" matrix, whose columns represent the time series associated with each component; $\boldsymbol{S}_{qxV}$ is the matrix of source components, whose rows represent statistically independent source signal maps; $\boldsymbol{E}_{TxV}$ is the error matrix that represents any variability not explained by the components. T and V represent the number of time points (scans) and the number of voxels, respectively.

ICA decomposes the fMRI data into a set of components, such that they exhibit a high degree of both within-component homogeneity, and between-component

Figure 1.6: ICA for fMRI data



(source: http://users.ics.aalto.fi/whyj/publications/thesis/thesis_node8.html)

variation. A major advantage of ICA is that it does not require any *a priori* assumptions about the spatiotemporal structure of the data, which is especially useful for analyzing resting-state data, where there are no clear task-related activations. ICA has the additional benefit of simultaneously separating neuronal and non-neuronal (e.g. head motion, respiration) sources into different components. Each estimated component represents a functional network, consisting of brain regions with a high degree of functional connectivity. Many studies have applied ICA to resting state fMRI to identify resting state networks (Laird et al., 2011; Smith et al., 2009).

ICA for fMRI has been well-documented in the single-subject case (McKeown et al., 1998; Beckmann and Smith, 2004). There are two primary approaches for extending ICA for multi-subject fMRI data: the Tensor PICA method (Beckmann and Smith, 2005) and the GIFT (Group ICA for fMRI toolbox) method (Calhoun et al., 2001). Guo (2011) propose a general statistical model for probabilistic ICA and a hierarchical model for group PICA (Guo and Tang, 2013). We use the well-established GIFT method, which can be easily implemented with a MATLAB toolbox.

The GIFT method can be summarized into the following steps:

1. Initial data reduction with PCA for each subject.

2. Temporal concatenation of the reduced data across subjects.

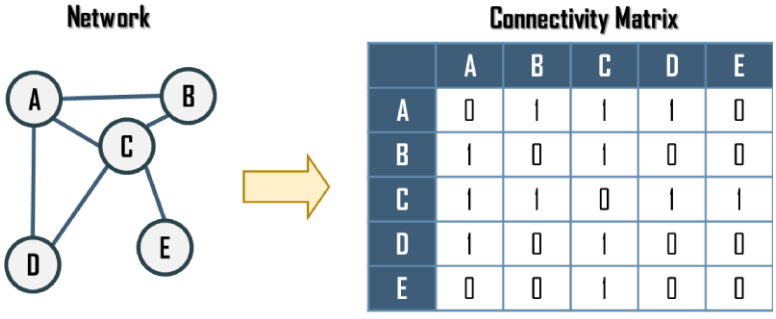
3. ICA decomposition of the concatenated data.

4. Back-reconstruction for subject-specific maps and time courses.

1.3.4.2 Network Modeling Methods for FC

An increasingly popular technique for modeling FC across the whole brain is through the network analysis approach. This method, borrowed from graph theory, models brain regions and the connections between them as a system of ”nodes” and ”edges”. The whole system can be represented by a graph or connectivity matrix (see Figure 1.7). This approach offer a way to visualize the functional connectivity network across the whole brain, and quantitatively characterize its properties with metrics like small-worldness, modularity, etc. (Bullmore and Sporns, 2009).

Figure 1.7: A simple network example



(source: https://people.hofstra.edu/geotrans/eng/methods/connectivitymatrix.html)

A typical functional neuroimaging network analysis begins by parcellating the brain into regions, then quantifying a measure of association between all pairs of regions, to produce a functional connectivity matrix (Rubinov and Sporns, 2010). The functional connectivity between a pair of regions is often measured by taking

the Pearson's correlation or partial correlation between the regions' fMRI time series (Smith et al., 2011). Partial correlation provides a more accurate estimate of the true network, because it regresses out the effects of all other nodes in the network before estimating the pairwise correlation. Compared to correlations, partial correlations are more computationally difficult to calculate for large networks, but Wang et al. (2016) describe a method for efficient and reliable estimation of the partial correlation matrix.

1.4 Structural Neuroimaging

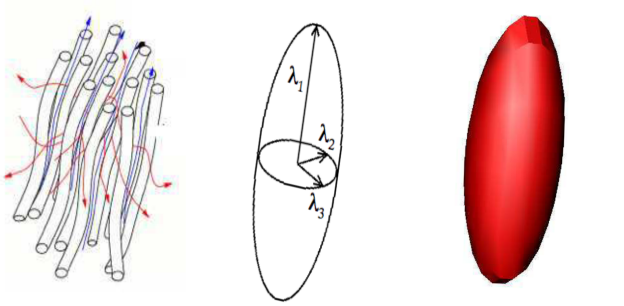
Unlike functional imaging modalities, which aim to measure changes in brain activity over time, the goal of structural imaging modalities is to examine brain anatomy and/or structural connectivity. Modalities such as T1-weighted MRI and CT are useful for obtaining high-resolution images of brain anatomy, while diffusion-weighted MRI (e.g. DTI) can be used to locate and quantify the structural connections (i.e. white matter fiber tracts) that physically connect remote brain regions.

1.4.1 Diffusion Tensor Imaging (DTI) Data

Diffusion Tensor Imaging (DTI) is a diffusion-weighted MRI technique that measures the direction and magnitude of water diffusion in brain tissue, allowing for the reconstruction of white matter fiber tracts via a tractography algorithm. Because of the fatty myelin sheath coating neuronal axon fibers, water diffusion is less restricted along the direction of fiber tracts, rather than across them. During a DTI scan, magnetic gradients are applied in different directions to enable calculation of the diffusion tensor at each voxel location (see Figure 1.8). The tensor is characterized by it's three principal eigenvectors and their associated eigenval-

ues (λ_1,λ_2,λ_3), and the shape of the estimated tensor indicates the directional dependency of the diffusion signal (Johansen-Berg and Rushworth, 2009). The level of diffusion restriction can captured by the Fractional Anisotropy (FA) measure, which ranges from 0 to 1. A completely spherical, isotropic tensor would have an FA value of 0, while an extremely elongated, anisotropic tensor would have an FA value close to 1.

Figure 1.8: Estimation of the diffusion tensor (DT) at a single voxel location



(Johansen-Berg and Rushworth, 2009)

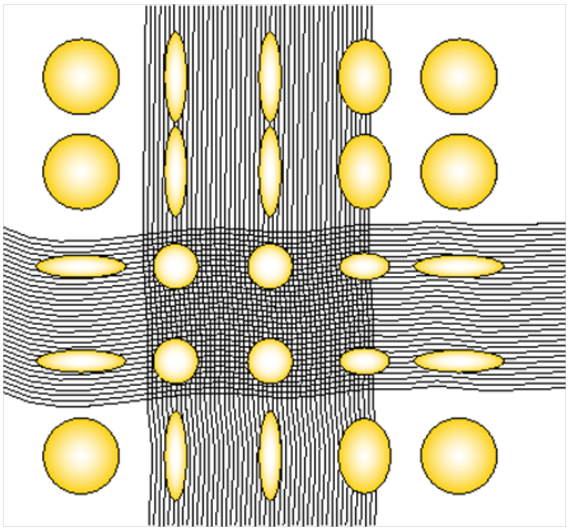
Figure 1.9 shows an example of two crossing fiber tracts (in black) and the shape of the tensors at each voxel (in yellow) within a small slice of the brain. Tensors outside of white matter tend to be spherical, indicating that diffusion is equally likely in all directions, while tensor inside a white matter fiber tract tend to be elongated to reflect the preferred direction of diffusion. Where white matter fibers cross, the shape of the tensor reflects that there is more than one preferred direction of diffusion.

1.4.2 DTI Preprocessing Pipeline

As with fMRI, several preprocessing steps are needed to prepare the DTI data for further analysis:

1. **Brain Extraction** to remove the skull and non-brain matter.

Figure 1.9: DTI example with two crossing fiber tracts



(source: http://www.rsierra.com/DA/node6.html)

2. **Eddy Current Correction** to correct for distortions and head movement.

3. **Registration** to standard coordinate space.

4. **Fit diffusion tensors** at each voxel.

5. **Fit a probabilistic diffusion model** on the corrected data. This step runs Markov Chain Monte Carlo sampling to build up distributions on diffusion parameters at each voxel.

6. **Run probabilistic (or deterministic) tractography** to determine the structural connectivity distribution from user-defined seed voxels or regions.

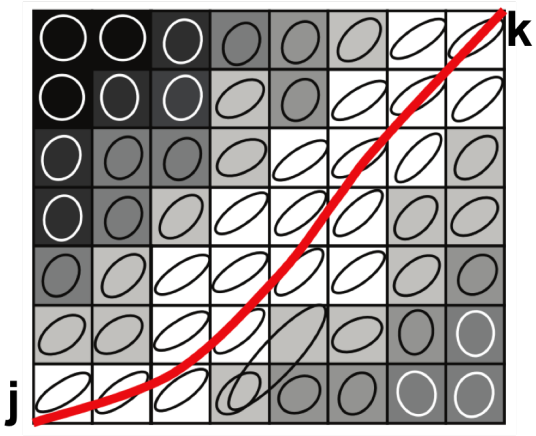
These preprocessing steps can be easily implemented with FSL's diffusion toolbox (FDT), among others (Jenkinson et al., 2012; Smith et al., 2004).

1.4.3 Tractography and Structural Connectivity (SC)

A tractography algorithm can determine the directional pattern of diffusion tensors to generate connectivity distributions from user-defined seed voxels. In a

bootstrap probabilistic tractography procedure, several streams are initialized at a seed voxel, and then traced through the diffusion tensor field to reconstruct and probabilistically quantify the fiber tracts in the brain. As illustrated in Figure 1.10, we can estimate the probability of structural connectivity between voxels j and k by $\hat{p}_{jk} = N_{jk}/N$, where N is the number of streams initialized from voxel j, and N_{jk} is the number of those streams that pass through voxel k. The procedure, implemented using FSL's FDT, allows for multiple fiber orientations and crossing fiber tracts (Behrens et al., 2007, 2003). An alternative streamline-tracking approach is deterministic tractography, which quantifies the strength of structural connectivity between a pair of brain locations using fiber counts (Gong et al., 2009).

Figure 1.10: probabilistic tractography



(Johansen-Berg and Rushworth, 2009)

Tractography algorithms provide a method of measuring the **structural connectivity (SC)** between a pair of brain regions. As with functional connectivity, SC can also be characterized across the whole brain using a network modeling approach (Bullmore and Sporns, 2009; Hagmann et al., 2007, 2008).

1.5 Combining Structure and Function

1.5.1 Motivation

Emerging evidence suggests that structural connections mediate functional connectivity, by providing a physical mechanism of communication between spatially remote brain areas (Collin et al., 2014; Hagmann et al., 2008; Van Den Heuvel et al., 2009; Toosy et al., 2004), although FC can still exist between brain regions that lack direct structural links (Honey et al., 2009; Damoiseaux and Greicius, 2009). Multimodal approaches that incorporate information from both fMRI and DTI can improve our understanding of the structure-function relationship in the brain, and allow us to characterize how these connections are affected by brain diseases.

Many brain disorders are associated with changes in connectivity. For example, Multiple Sclerosis patients exhibit white matter atrophy, in which the myelin sheath surrounding axons is destroyed, disrupting communication between neurons. Alzheimer's Disease is also associated with disruptions of SC and FC (Filippi and Agosta, 2011; Bozzali et al., 2011), as are many other brain disorders (Zhang and Raichle, 2010; Greicius, 2008). By integrating structural and functional information, we can more comprehensively characterize the pathophysiology of brain diseases.

1.5.2 Review of Existing Multimodal Methods

Multimodal neuroimaging methods combine information from two or more datasets acquired by different imaging techniques, each of which provides a limited view of the brain. The goal of this approach is to capitalize on the strength of each modality in a joint analysis, rather than separately analyzing each modality. For this dissertation, we will review methods that combine information from functional and structural neuroimaging modalities, namely fMRI and DTI.

There have been several recent attempts to examine the FC-SC relationship by combining information from fMRI and DTI data (Rykhlevskaia et al., 2008; Damoiseaux and Greicius, 2009). Many of these studies perform analyses on each modality separately, then assess the concordance between the FC and SC measures (Honey et al., 2009; Greicius et al., 2009; Skudlarski et al., 2008; Rudie et al., 2013; Horn et al., 2013). Others have developed brain connectivity methods that combine information from functional and structural modalities. Bowman et al. (2012), for example, present a framework for an anatomically-weighted FC measure that integrates fMRI and DTI data, and demonstrate that the supplemental SC information is particularly beneficial in the presence of fMRI noise. Also, recent multimodality fusion approaches such as linked ICA (Calhoun et al., 2006), multimodal canonical correlation analysis (Correa and Li, 2009), and joint ICA (Groves et al., 2011), along with joint models of the FC and SC networks (Xue et al., 2015; Venkataraman et al., 2012) have shown that a joint analysis of data from both modalities performs better than separate analyses of FC or SC alone. A review of multimodal fusion methods is provided by Sui et al. (2012) and Zhu et al. (2013).

Despite the evidence of structure-function associations in the brain, the functional connectivity methods described in section 1.3.4 fail to incorporate information about the underlying SC. Including structural data in these analyses could allow researchers to create more biologically plausible models of functional connectivity in the brain. As the relationship between structure and function unfolds, an important next step is to develop statistical methods that combine modalities and allow a convenient framework to conduct inference.

In addition, network approaches have become increasingly useful for modeling the complex patterns of functional and structural connectivity across the whole-brain network. Network studies of psychopathology rely on the notion that dysfunctional nodes and edges result in abnormal signaling, which results in differ-

ences in brain connectivity networks. Since most major brain disorders involve dysfunction of processes that are distributed across multiple brain regions, rather than being constrained to one individual brain area, the network approach can provide important insights on global brain connectivity and organization in psychopathology (Menon, 2011).

1.6 Motivating Data Example

For some brain disorders, a clear disruption in brain connectivity has been established and corroborated by results from multiple studies. In Alzheimer's Disease for example, studies have consistently found decreased connectivity in the default mode network and the hippocampus (Zhang and Raichle, 2010). However, the mechanism of connectivity disruption is not as well defined in all neurological and psychiatric disorders.

1.6.1 Major Depressive Disorder (MDD)

Major depressive disorder (MDD) affects approximately 1 in 6 people in their lifetimes, causes substantial occupational and social impairments, and is comorbid with anxiety and substance abuse disorders (Kessler et al., 2003). Although there have been many resting-state studies conducted on MDD, they do not all agree about connectivity, and its pathology is still unclear (Northoff et al., 2011). (Anand et al., 2009) found that untreated MDD patients exhibited decreased connectivity between the dorsal ACC and other areas, and that connectivity improved in the disrupted pathways after 6 weeks of treatment. Greicius et al. (2007) showed that MDD patients had increased connectivity between certain regions of the default mode network, and De Kwaasteniet et al. (2013) suggests that changes in the FC-SC relationship play a role in MDD pathophysiology. Further studies of functional

and structural neuroimaging play critical roles in advancing our knowledge about this disorder.

1.6.2 Subjects

The data used in this dissertation was collected by the Mayberg Lab in the Emory Department of Psychiatry. The dataset consists of 20 patients with MDD, and 20 healthy controls, matched by age and gender. The average ages of the MDD and control groups were 45.8 and 42.4 years, respectively; both groups were 50% male. For the MDD patients, the average Hamilton Depression Rating Scale (HAM-D) score was 19, indicating "Severe Depression" (Hamilton, 1960), and the average length of current episode was 82 weeks.

The dataset includes resting-state fMRI and DTI scans for each subject at baseline. At the time that the scans were collected, all MDD patients had unipolar depression and had never been treated. During the scan, subjects were instructed to lie passively in the scanner and to refrain from thinking about anything in particular. Since resting-state captures the brain "in idle", and the brains of MDD patients appear to be "stuck" in an idle state, the resting-state paradigm is well-suited for studying this disorder.

1.6.3 Data acquisition and preprocessing

DTI, rs-fMRI, and T1-weighted MRI data were collected in a single session with 3T Siemens Tim Trio scanner. The DTI sequence consisted of 60 scans with different diffusion-weighted directions ($b = 1000s/mm^2$) and four non-diffusion weighted scans ($b = 0$), acquired using a single-shot spin-echo echo planar imaging (EPI) sequence. Additional DTI scanning parameters include: TR=11300 ms, TE=104 ms, GRAPPA on, FOV=256 mm, number of slices=64, voxel size=2x2x2 mm, and

matrix size=128x128. For registration purposes, high-resolution T1-weighted images were collected using a 3D MPRAGE sequence with the following parameters: TR=2600 ms, TI=1100 ms, TE=3 ms, number of slices = 176, voxel size = 111 mm, matrix size= 224256, flip angle=8°. Functional images were collected over 150 time points, with a z-saga sequence to minimize artifacts in the medial PFC and OFC due to sinus cavities. Scans were acquired interleaved with the following parameters: TR=2.92 seconds; TE1=30 ms, TE2=66 ms, flip angle=90°, number of axial slices=30, slice thickness=4 mm, FOV=220 mm, and total duration=7.3 min.

Several standard preprocessing steps were applied to the functional images, including despiking, slice timing correction, motion correction, registration to MNI 2mm standard coordinate space, normalization to percent signal change, removal of linear trend, regressing out CSF, WM, and 6 movement parameters, bandpass filtering (0.009 to 0.08), and spatial smoothing with a 6mm FWHM Gaussian kernel. Preprocessing steps for the DTI data include brain extraction to remove non-brain regions, phase reversal distortion correction, and aligning diffusion weighted images to the average non-diffusion weighted image by rigid body affine transformation to remove motion and eddy-current induced artifact. Using the Diffusion Toolbox (FDT) in FSL, we can then fit a diffusion tensor model at each voxel, to quantify the directional diffusion at that brain location (Behrens et al., 2003).

1.7 Proposed Research

The main objective of this dissertation is to develop statistical methods for multimodal neuroimaging data that allow us to explore the relationship between brain FC and SC, and investigate how this relationship varies across different kinds of brain connections, in order to distinguish different types of FC, e.g. FC based on direct SC vs. FC due to global effects or functional co-activations. In addition,

we aim to explore potential between-group differences in the FC-SC association, specifically, comparing MDD subjects vs. healthy controls.

1.7.1 Topic 1: Quantifying the strength of structural connectivity underlying functional brain networks

For the first topic, we propose a novel measure to quantify the strength of structural connectivity (sSC) underlying functional networks, by incorporating structural information from DTI data. To conduct inference on our sSC measure, we estimate a covariance term that considers spatial similarity between observations via a parametric semivariogram model with a novel distance metric. We demonstrate the performance of our proposed measure with simulation studies, and apply our method to the fMRI and DTI dataset described in section 1.6. Since some functional networks are more reproducible across different studies/subjects, we explore whether sSC is associated with component reliability, and demonstrate the benefit of leveraging information from structural data in the estimation of functional networks from fMRI data.

1.7.2 Topic 2: A joint model for functional and structural connectivity across the whole-brain network

The second and third topics propose statistical frameworks for modeling the relationship between FC and SC across the whole-brain network. The second topic presents a hierarchical model with a linear link function to describe the association at each edge in the network, and uses the EM algorithm to estimate the model parameters. We consider both correlation and partial correlation as potential measures of functional connectivity. We conduct simulation studies to evaluate performance of our model, and apply the proposed methods to the fMRI and DTI

dataset, demonstrating biologically meaningful findings.

1.7.3 Topic 3: Using copulas to model the structure-function relationship in the brain

As an extension to topic 2, the third topic considers using copulas to construct a more flexible and robust statistical measure of the FC-SC association. The copula approach allows us to flexibly model the dependence structure between FC and SC, in a way that does not depend on the marginal distributions of the connectivity values. Simulation studies and data analysis demonstrate the utility of this method.

Chapter 2

Topic 1: Quantifying the strength of structural connectivity underlying functional brain networks

2.1 Introduction

Independent component analysis (ICA), which was reviewed in section 1.3.4.1, is a popular approach for estimating functional networks from fMRI data (Beckmann and Smith, 2004; Calhoun et al., 2001; Guo and Pagnoni, 2008). In addition, we can use DTI data to extract information about the underlying structural connections (SC) that physically connect remote brain regions. Since many studies have shown evidence that SC mediates functional connectivity in the brain (Greicius et al., 2009; Van Den Heuvel et al., 2009), data-driven FC methods like ICA could be improved by incorporating information about the underlying SC. Another limitation of partitioning algorithms like ICA is that the reliability of the estimated functional networks is unknown. For example, the results from an ICA run may vary based on the choice of algorithm type, algorithm starting points, subject variability, or data preprocessing steps (Calhoun et al., 2004). Therefore, some functional networks may not be reproducible in different analyses or in other data sets. Determining the reliability of the estimated ICs obtained from a given data set is important for making correct interpretations. Previous work has proposed assessing reliability of ICs by evaluating the reproducibility in repeated ICA runs, with different initial conditions, with data resampled from the original functional time series, or with simulated data having a known brain network structure (Himberg et al., 2004; Meinecke et al., 2002; Duann et al., 2006). These approaches measure the reliability of ICs only through the algorithmic and statistical reproducibility of ICA applied to fMRI data. The advent of DTI provides an opportunity to leverage information about SC to take a more sound and reliable approach to estimating functional networks from fMRI data.

As the relationship between brain structure and function unfolds, it is important to develop statistical methods that combine information from fMRI and DTI data. We present a novel measure of the strength of structural connectivity (sSC)

underlying a functional network. We aim to use this measure to characterize the sSC underlying FC networks estimated by ICA, develop a statistical testing framework for the measure, and compare sSC between subject groups. We also propose a novel reliability index for estimated FC networks, based on bootstrapped ICA runs, and investigate the association between an FC network's underlying sSC and its reliability. We conduct simulation studies to evaluate the performance of our proposed measure, and apply our method to a resting-state fMRI and DTI dataset.

2.2 Data

We use the dataset described in section 1.6, that contains resting-state fMRI (rs-fMRI) and DTI scans for 20 subjects with major depressive disorder (MDD) and 20 healthy controls.

2.2.1 Identifying Functional Networks

We identify group-level FC network maps by running a spatial ICA on the control subjects' rs-fMRI data via the Group ICA for fMRI Toolbox (GIFT) in Matlab (Calhoun et al., 2001). We use only the control subjects' fMRI data to generate a common set of FC network maps, since MDD is known to be associated with functional abnormalities in resting-state networks (Veer, 2010; Greicius et al., 2007; Northoff et al., 2011). We apply a binary coverage mask, created from the intersection of all subjects' brain masks, to ensure that only voxels within the brain are used in the analysis. The group ICA yields a common set of spatially independent component (IC) maps and their associated time series. Many of the IC maps correspond to known resting state networks, such as the default mode network (Buckner et al., 2008), while others represent noise and artifacts. To determine which ICs reflect

true FC networks, we evaluate their correlation with the set of major resting-state network maps defined by Smith et al. (2009); we use the cutoff of $r \geq 0.25$ to determine whether a IC corresponds to a known network.

2.2.2 Determining Structural Connectivity

We can evaluate structural connectivity across the whole brain from DTI data by implementing a widely-used probabilistic tractography approach via FSL's diffusion toolbox (Behrens et al., 2007, 2003). We identify the set of white matter voxels in each ICA map estimated from the group ICA, to be used as seed locations in the probabilistic tractography procedure. In this procedure, we successively initiate a given number of streams (5000 in our analysis) from each seed voxel, and trace the streams as they pass through the diffusion tensor field, following the principle direction of diffusion from voxel to voxel. The streams are constrained to remain in white matter and avoid cerebral spinal fluid (CSF), and terminate according to a stopping rule. Thus, the number of streams (out of 5000 trials) that extend from the seed voxel to a target voxel empirically quantifies the probability of SC between these two brain locations.

2.3 Methods

2.3.1 The strength of Structural Connectivity (sSC) measure

Suppose we are investigating q IC maps extracted from a group ICA of fMRI data. To quantify the strength of structural connectivity (sSC) underlying the FC net-

work specified by the ℓ^{th} IC ($\ell = 1, ..., q$), we propose the following measure, θ_ℓ:

$$\theta_\ell = \frac{\sum_{j,k \in \Omega_\ell} [p_{jk} - (\bar{p_j} + \bar{p_k})/2]}{\sum_{j,k \in \Omega_\ell} [1 - (\bar{p_j} + \bar{p_k})/2]}$$

Here, Ω_ℓ is the set of voxels within the ℓ^{th} IC; p_{jk} is the probability of SC between the pair of voxels j and k within IC ℓ; $\bar{p_j}$ is the average probability of connection between voxel j and the rest of the brain, defined as

$$\bar{p_j} = \frac{1}{V-1} \sum_{\substack{v=1 \\ v \neq j}}^{V} p_{jv}$$

where V is the total number of voxels in the brain. ($\bar{p_k}$ is defined analogously.)

The second term in the numerator, $(\bar{p_j} + \bar{p_k})/2$, represents the average overall probability of SC between the voxels j, k in IC ℓ and a random location in the brain. Thus, the full numerator of θ_ℓ reflects the degree to which the actual SC within the FC network exceeds the SC between the network and the rest of the brain expected on average. This adjustment allows us to compare ICs with different levels of overall connectedness. We standardize the sSC measure by dividing by the maximum possible value, in which there is complete SC between all voxel pairs within the FC network (i.e. $p_{jk} = 1$ for all pairs j, k). This standardization will restrict θ_ℓ to be less than or equal to 1, and allow comparison of this measure across FC networks of different sizes. In this way, the structure of θ_ℓ resembles that of the Kappa measure of inter-rater agreement, as it represents the observed underlying strength of SC, relative to the SC expected by chance, divided by the maximum possible value.

We can estimate the probability of SC between any pair of brain locations using a probabilistic tractography procedure with DTI data, as described in section 2.2.2.

In this procedure, we initiate N streams from seed voxel j and track how many of these streams pass through target voxel k; the number of streams connecting these voxels is denoted by N_{jk}. Thus, we can estimate p_{jk}, the probability of SC for voxel pair j, k, by $\hat{p}_{jk} = N_{jk}/N$.

In this way, we can use the results of the probabilistic tractography procedure to construct the following estimate of θ_ℓ:

$$\hat{\theta}_\ell = \frac{\sum_{j,k \in \Omega_\ell} [N_{jk} - (\bar{N}_j + \bar{N}_k)/2]}{\sum_{j,k \in \Omega_\ell} [N - (\bar{N}_j + \bar{N}_k)/2]}$$

where N_{jk} is the number of streams passing through voxels j and k, $\bar{N}_j$ and $\bar{N}_k$ are the average N_{jv} and N_{kv} over all voxels v in the brain, respectively, and N is the total number of streams initiated in the procedure. Thus, the $\hat{\theta}_\ell$ statistic represents the above-average strength of SC underlying an FC network estimated by ICA.

Although we've defined $\hat{\theta}_\ell$ here on the voxel-level, this measure could also be defined on the region-level, where j and k represent a pair of regions rather than voxels. When performing probabilistic tractography on the voxel-level, the magnitude of $\hat{\theta}_\ell$ will inherently be small relative to its upper bound of 1, because the probability of SC between two individual voxels tends to be very low due to the small target size. However, the magnitude of $\hat{\theta}_\ell$ based on region-to-region probabilistic tractography is expected to be larger because each region contains multiple voxels.

To conduct inference for θ_ℓ, we will need an estimate of its variance, $var(\hat{\theta}_\ell)$.

First, consider $\boldsymbol{N^*}$, a $\binom{V}{2} \times 1$ vector of N_{jk} for the set of all voxel pairs $\{j, k\}$.

$$\boldsymbol{N^*} = \Big[\{N_{jk}\}\Big] = \begin{bmatrix} N_{12} \\ N_{13} \\ \vdots \\ N_{V-1,V} \end{bmatrix}$$

Our proposed strength of SC statistic, $\hat{\theta}_\ell$, can then be written as a function of $\boldsymbol{N^*}$ (see Appendix A for details):

$$\hat{\theta}_\ell = \frac{(\boldsymbol{C_\ell} - \boldsymbol{A})\boldsymbol{N^*}}{b - \boldsymbol{A}\boldsymbol{N^*}}$$

where

$$\boldsymbol{A} = \frac{(V_\ell - 1)}{2(V-1)} \sum_{j \in \Omega_\ell} \boldsymbol{C_j}$$

and

$$b = \frac{V_\ell(V_\ell - 1)}{2} N$$

and $\boldsymbol{C_\ell}$ and $\boldsymbol{C_j}$ are $1 \times \binom{V}{2}$ row vectors of binary indicators, corresponding to the order of the $\{N_{jk}\}$ in $\boldsymbol{N^*}$. The elements of $\boldsymbol{C_\ell}$ are 1 if the j, k voxel pair is a member of Ω_ℓ, 0 otherwise; the elements of $\boldsymbol{C_j}$ are 1 if the j, k voxel pair involves voxel j, 0 otherwise. V is the total number of voxels in the whole brain and V_ℓ is the number of voxels in IC ℓ.

We can consider each $N_{jk} \sim \text{Bin}(N, p_{jk})$ where N is the number of streams initialized in a probabilistic tractography procedure, and p_{jk} is the probability of SC between voxels j and k. Given that N is large, we can approximate with the Normal distribution, $N_{jk} \sim \text{N}(\mu_{jk}, \sigma^2_{jk})$, and $\boldsymbol{N^*}$ can be considered to follow a multivariate normal distribution with mean vector $\boldsymbol{\mu}$ of size $\binom{V}{2} \times 1$ and variance-covariance matrix $\boldsymbol{\Sigma}$ of size $\binom{V}{2} \times \binom{V}{2}$.

That is, $\boldsymbol{N^*} \sim \text{MVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where

$$\boldsymbol{\mu} = N\left[\{p_{jk}\}\right] = N\begin{bmatrix} p_{12} \\ p_{13} \\ \vdots \\ p_{V-1,V} \end{bmatrix}$$

and

$$\begin{aligned} \boldsymbol{\Sigma} &= \left[\{cov(N_{jk}, N_{j'k'})\}\right] \\ &= \begin{bmatrix} var(N_{12}) & cov(N_{12}, N_{13}) & \cdots & cov(N_{12}, N_{V-1,V}) \\ cov(N_{13}, N_{12}) & var(N_{13}) & \cdots & cov(N_{13}, N_{V-1,V}) \\ \vdots & \vdots & \ddots & \vdots \\ cov(N_{V-1,V}, N_{12}) & cov(N_{V-1,V}, N_{13}) & \cdots & var(N_{V-1,V}) \end{bmatrix} \end{aligned}$$

Because the numerator and denominator of $\hat{\theta}_\ell$ are linear combinations of $\boldsymbol{N^*}$, we can use the Delta method based on $\boldsymbol{\Sigma}$ to derive the approximate the large sample variance of $\hat{\theta}_\ell$ (Casella and Berger, 1990).

$$\begin{aligned} var(\hat{\theta}_\ell) &= var\left(\frac{(\boldsymbol{C_\ell} - \boldsymbol{A})\boldsymbol{N^*}}{b - \boldsymbol{A}\boldsymbol{N^*}}\right) \\ &\approx \left[\frac{(\boldsymbol{C_\ell} - \boldsymbol{A})\boldsymbol{\mu}}{b - \boldsymbol{A}\boldsymbol{\mu}}\right]^2 \left[\frac{(\boldsymbol{C_\ell} - \boldsymbol{A})\boldsymbol{\Sigma}(\boldsymbol{C_\ell} - \boldsymbol{A})'}{[(\boldsymbol{C_\ell} - \boldsymbol{A})\boldsymbol{\mu}]^2} + \frac{\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A'}}{[b - \boldsymbol{A}\boldsymbol{\mu}]^2} - 2\frac{[-(\boldsymbol{C_\ell} - \boldsymbol{A})\boldsymbol{\Sigma}\boldsymbol{A'}}{[(\boldsymbol{C_\ell} - \boldsymbol{A})\boldsymbol{\mu}][b - \boldsymbol{A}\boldsymbol{\mu}]}\right] \end{aligned}$$

Although $var(\hat{\theta}_\ell)$ is a scalar value, its calculation requires estimation of the large $\binom{V}{2} \times \binom{V}{2}$ matrix $\boldsymbol{\Sigma}$. Estimating $\hat{\boldsymbol{\Sigma}}$ poses computational difficulties because it is a high-dimensional matrix with spatially dependent observations (i.e. the probability of SC tends to be higher for brain locations that are spatially close). We can incorporate the spatial distance between observations in our estimation of $\hat{\boldsymbol{\Sigma}}$ by

using a parametric semivariogram model (Minasny and McBratney, 2005). In this way, we can model each covariance term, $cov(N_{jk}, N_{j'k'})$, as a function of the distance between voxel pair j, k and voxel pair j', k', denoted $d_{jk,j'k'}$, which decays as the distance between the observations increases. The parameters of the semivariogram can be estimated from the data, by fitting to the empirical semivariogram function.

The elements in $\boldsymbol{N}^*$ are based on observations between voxel pairs, so to capture the spatial similarity between these observations in the covariance term, we need a metric that quantifies the spatial distance between *pairs* of voxel pairs j, k and j', k'. Since this is a non-standard distance, we propose a novel distance metric to incorporate in our spatial model for covariance:

$$d_{jk,j'k'} = \min\left[\frac{d(j,j') + d(k,k')}{2}, \frac{d(j,k') + d(k,j')}{2}\right]$$

where $d(j,k)$ is the euclidean distance between voxels j and k; this serves as a proxy for the true distance between voxels that is based on the length of the fiber tract. Thus, we can use this distance metric and the semivariogram model to estimate the elements of $\hat{\boldsymbol{\Sigma}}$, then calculate $var(\hat{\theta}_\ell)$ to use in hypothesis testing for θ_ℓ. To avoid the computational burden of estimating the large matrix $\hat{\boldsymbol{\Sigma}}$, we will also consider a bootstrap standard error term, denoted $SE_{boot}(\hat{\theta}_\ell)$, as a non-parametric alternative for the theoretical variance term, $var(\hat{\theta}_\ell)$. We will compare the performance of the theoretical vs. bootstrap variance terms via simulation studies in section 2.4.

2.3.2 Hypothesis testing based on the sSC measure

Our goal in developing the strength of SC measure, $\hat{\theta}_\ell$, and its variance term is to evaluate the strength of SC underlying functional networks estimated by data-

driven partitioning methods like ICA. We propose a hypothesis testing framework for 1) evaluating sSC within a single independent component (IC), 2) comparing sSC between two ICs, and 3) comparing the sSC for a given IC between two subject subgroups (i.e. MDD vs controls).

For an individual IC ℓ, if $\theta_\ell = 0$, this indicates that observed strength of SC within the functional network is no higher than we would expect based on its average SC. However, if an IC is representative of a true functional network, we expect the strength of SC underlying these functional connections to be significantly *above* average, i.e. $\theta_\ell > 0$. Thus, we can evaluate the strength of SC within IC ℓ using the hypotheses:

$$H_0 : \theta_\ell = 0 \text{ vs. } H_a : \theta_\ell > 0$$

and we construct a Wald-type test statistic for the single- and multi-subject cases, respectively:

$$T^* = \frac{\hat{\theta}_\ell}{\sqrt{\hat{Var}(\hat{\theta}_\ell)}} \sim N(0,1) \qquad T^* = \frac{\bar{\hat{\theta}}_\ell}{\sqrt{\hat{Var}(\hat{\theta}_\ell)/n}} \sim N(0,1)$$

In order to compare the strength of SC between two ICs, ℓ and ℓ', we can use the hypotheses:

$$H_0 : \theta_\ell = \theta_{\ell'} \text{ vs. } H_a : \theta_\ell \neq \theta_{\ell'}$$

and evaluate using a non-parametric permutation testing approach, in which we permute the IC label to generate an empirical distribution for the $\bar{\hat{\theta}}_\ell - \bar{\hat{\theta}}_{\ell'}$ difference.

Finally, we can test whether the strength of SC for a given IC ℓ differs between subject subgroups (i.e. MDD vs controls) using the hypotheses:

$$H_0 : \theta_{\ell,1} = \theta_{\ell,2} \text{ vs. } H_a : \theta_{\ell,1} \neq \theta_{\ell,2}$$

and a Wald-type test statistic:

$$T^* = \frac{\bar{\hat{\theta}}_{\ell,1} - \bar{\hat{\theta}}_{\ell,2}}{\sqrt{\frac{\hat{Var}(\hat{\theta}_{\ell,1})}{n_1} + \frac{\hat{Var}(\hat{\theta}_{\ell,2})}{n_2}}} \sim N(0,1)$$

Alternatively, we can test the between-group hypotheses using a non-parametric permutation testing approach, in which we permute subject group label to generate an empirical distribution for the $\bar{\hat{\theta}}_{\ell,1} - \bar{\hat{\theta}}_{\ell,2}$ difference.

2.3.3 Using sSC to inform reliability of components

Due to the stochastic nature of the ICA algorithm and individual subject variability, the reliability of the functional networks estimated from fMRI data may vary. However, it is important to determine the reliability of the estimated ICs to make accurate interpretations. To investigate whether strength of SC can be leveraged to inform IC reproducibility, we propose a novel reliability index based on bootstrapped ICA runs.

To calculate this reliability index, we first generate B bootstrap samples of size n and perform group ICA for each bootstrap sample, extracting q ICs each time. Next, for IC ℓ from the original data, we identify the corresponding IC ℓ_b in the b^{th} bootstrap sample. We then calculate $r_{\ell\ell_b}$, the spatial correlation between original IC ℓ and its corresponding bootstrap IC ℓ_b, as well as $r_{\ell\ell_b^*}$, the spatial correlation between IC ℓ and each other IC ℓ_b^* in the b^{th} bootstrap sample. Based on these correlation values, we propose the following reliability index R_ℓ for IC ℓ:

$$R_\ell = \frac{\frac{1}{B}\sum_{b=1}^{B} |r_{\ell\ell_b}| - \frac{1}{Bq}\sum_{b=1}^{B}\sum_{\ell_b^*=1}^{q} |r_{\ell\ell_b^*}|}{1 - \frac{1}{Bq}\sum_{b=1}^{B}\sum_{\ell_b^*=1}^{q} |r_{\ell\ell_b^*}|}$$

The numerator of R_ℓ represents the observed similarity between IC ℓ and its corre-

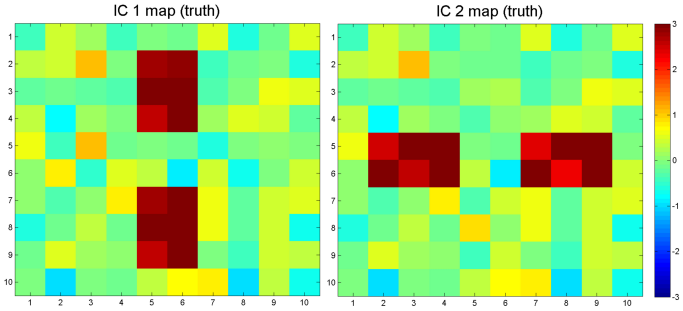
sponding bootstrapped IC ℓ_b, relative to its average similarity to all of the q bootstrapped ICs ℓ_b^*, and averaged over B bootstrap samples. We standardize the measure by dividing by the maximum possible value, in which original IC ℓ is perfectly reproduced in the bootstrap sample (i.e. $r_{\ell\ell_b} = 1$). The reliability index R_ℓ ranges from -1 to 1, where $R_\ell < 0$ indicates that the IC is not reliably reproduced across bootstrap samples, and R_ℓ close to 1 indicates that the IC is highly reproducible. Similar to the construction of the sSC measure θ_ℓ, R_ℓ is structured like the Kappa statistic for inter-rater agreement, in that it represents the observed reproducibility of IC ℓ, corrected for the reproducibility expected by chance, and divided by the maximum possible value.

2.4 Simulation Studies

We conducted simulation studies to evaluate the performance of the estimation and inference methods for the proposed strength of SC measure. We draw 300 simulated data sets for $n = 20, 50$ subjects, under two different noise levels. First, we simulate the fMRI data based on the true source signal maps and their time courses. We define $q = 2$ source IC maps, which are common for all subjects and consist of one 10×10 axial slice, for a total of $V = 100$ voxels. IC 1 represents a symmetric front-back network, while IC 2 represents a symmetric left-right network. The value at each voxel in these maps is based on the background noise ($x_1 \sim N(0, 0.5)$ for all voxels, plus the within-source intensity ($x_2 = 3$) and noise ($x_3 \sim N(0, 0.1)$) for voxels within the IC map. After the IC maps are created, their temporal responses are adapted from real fMRI data with $T = 200$ time points. We generate a $T \times V$ fMRI data matrix Y_i, for each subject i, according to the ICA equation: $Y_i = AS + e$, where A is the $T \times q$ "mixing matrix" whose columns represent the time series for each IC, and S is the $q \times V$ source matrix whose rows represent

the IC maps. Gaussian background noise is linearly added to the mixed spatial sources to generate a simulated fMRI data matrix of size 200×100 for each subject.

Figure 2.1: True IC source maps specified for simulation testing



Once the subject-level fMRI data has been simulated, we can estimate the group-level IC maps using the GIFT method (Calhoun et al., 2001). In order to evaluate the strength of SC underlying each estimated IC, we must first simulate the probabilistic tractography SC results based on DTI data. For each subject, we generate the $\binom{V}{2} \times 1$ matrix $\boldsymbol{N}^*$, whose elements (N_{jk}) represent the number of streams out of $N = 20$ trials that connect each voxel pair (j, k), to simulate the results of a probabilistic tractography procedure. We use the model $\boldsymbol{N}^* \sim \text{MVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ to simulate this data, with $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ defined as follows. For the mean vector $\boldsymbol{\mu} = N\boldsymbol{p}$, where $\boldsymbol{p}$ is the vector of voxel pair connection probabilities. Voxel pairs outside of an IC map have a connection probability of 0.25, while voxel pairs inside IC 1 or 2 have connection probabilities of 0.5 and 0.75, respectively. The covariance elements of the variance-covariance matrix $\boldsymbol{\Sigma}$ are defined based on the exponential semivariogram function,

$$\gamma(h) = c_0 + c_e[1 - e^{-|h|/a_e}], \qquad h > 0$$

with parameters c_0 (nugget), c_e (partial sill), and a_e (range); h represents the distance between observations. We generate the SC data in $\boldsymbol{N}^*$ under both "low" (c_0=1, c_e=4, a_e=1) and "high" (c_0=2, c_e=5, a_e=1) noise conditions. In this way, we can simulate the SC data in $\boldsymbol{N}^*$, and evaluate the strength of SC underlying each

IC.

Table 2.1 summarizes the results based on 300 simulation runs under the four different sampling configurations. In each setting, we estimate the strength of SC measure, $\hat{\theta}_1$ and $\hat{\theta}_2$, along with its variance and 95% confidence intervals, based on both the theoretical variance term and the bootstrap standard error (using B=1000 bootstrap resamples).

Table 2.1: Results based on 300 simulation runs

	n	Noise Level	θ	$\hat{\theta}$ mean (SD)	SE (Theoretical)	SE (Bootstrap)	Cov Prob (Theoretical)	Cov Prob (Bootstrap)
IC 1	20	Low	0.3077	0.3081 (0.0091)	0.0083	0.0093	92.6	94.3
		High		0.3074 (0.0104)	0.0093	0.0105	91.6	94
	50	Low		0.3084 (0.0061)	0.0053	0.0060	90.7	93.7
		High		0.3078 (0.0069)	0.0059	0.0068	91	94.7
IC 2	20	Low	0.64	0.6405 (0.0120)	0.0112	0.0115	93.3	93
		High		0.6389 (0.0134)	0.0126	0.0127	94.6	93.6
	50	Low		0.6409 (0.0080)	0.0071	0.0074	90.7	93
		High		0.6394 (0.0088)	0.0080	0.0082	93.7	93

We evaluate the bias of our sSC estimator by comparing the mean $\hat{\theta}$ to the true sSC value, θ; there is very low bias in all simulation settings. We also assess the performance of our two candidate variance terms by comparing the estimated theoretical and bootstrap standard errors (SE) to the empirical SE, SD($\hat{\theta}$). We note that the theoretical SE (based on semivariogram model fitting for $\hat{\mathbf{\Sigma}}$) tends to underestimate the variability of θ, while the bootstrap SE performs fairly well. Finally, we compare coverage probabilities based on two types of 95% confidence intervals (CIs): the Wald-type CI based on the theoretical SE, and the CI based on the bootstrap percentiles. The coverage probabilities from both types of CIs are fairly close to 95%, although the bootstrap CI tends to outperform the theoretical variance-based CI.

Because it requires estimating the large matrix $\hat{\mathbf{\Sigma}}$, calculation of the theoretical variance term $Var(\hat{\theta})$ poses a computational challenge, especially when the number of voxels V is large. Even in this small-scale simulation study where $V = 100$,

$\hat{\Sigma}$ has dimensions $\binom{100}{2} \times \binom{100}{2}$, or 4950 $\times$ 4950. The bootstrap variance estimator, on the other hand, is more computationally feasible since it avoids estimation of $\hat{\Sigma}$, and shows good performance in our simulation studies. Thus, we recommend using the bootstrap method to conduct inference in real data applications where the number of voxels V is large.

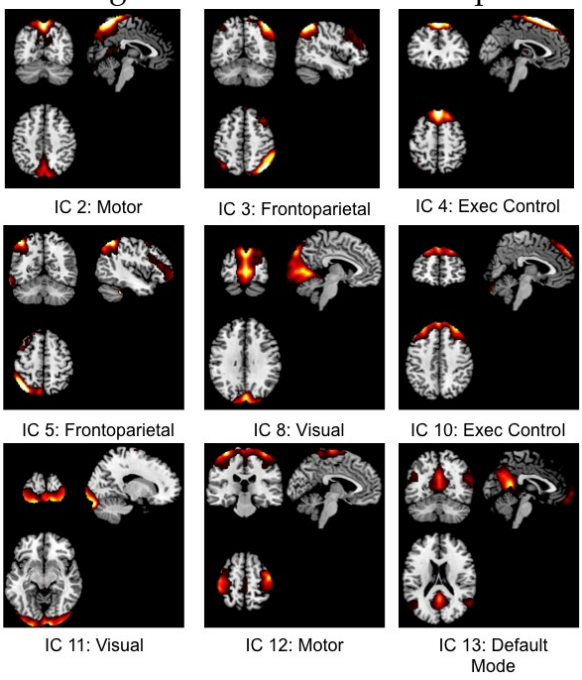
2.5 Data Analysis

We apply our sSC method to an fMRI and DTI dataset of 20 subjects with major depressive disorder (MDD) and 20 healthy controls. Initially, we run a group ICA using only the 20 control subjects' fMRI data, since studies of MDD have shown resting state FC differences (Northoff et al., 2011; Greicius et al., 2007; Veer, 2010). We extract $q = 15$ group-level IC maps, 9 of which appear to represent well-known resting state networks (Smith et al., 2009; Laird et al., 2009) (see Figure 2.2). We create a thresholded white matter mask for each IC, consisting of about 900 voxels, for further exploration of the underlying structural connectivity.

To evaluate the SC distribution of each IC, we run a probabilistic tractography procedure using the voxels in the thresholded IC mask as seed locations. We initiate N=5000 streams from each seed voxel in the IC, and trace the streams as they pass through the brain. The results of this procedure give us N_{jk}, $\bar{N}_j$, and $\bar{N}_k$ for each voxel pair j, k in IC ℓ, which can be used to estimate the strength of SC measure, $\hat{\theta}_\ell$. We conduct inference for θ_ℓ using the bootstrap SE term, rather than the theoretical $Var(\hat{\theta}_\ell)$ term for computational feasibility.

Table 2.2 shows the sSC results for the 9 estimated ICs from the control subject group. The $\hat{\theta}$ values are all fairly small, since our it was calculate on the voxel level, yet all ICs have strength of underlying SC significantly greater than 0. Since these IC maps all correspond to known resting state functional networks, it is not sur-

Figure 2.2: Estimated IC maps



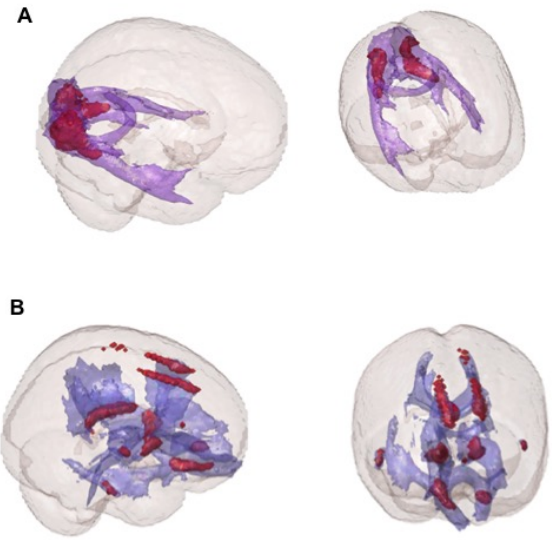
prising that they demonstrate within-network strength of SC above baseline. IC 8 displays the highest mean $\hat{\theta}$, while IC 4 displays the lowest mean $\hat{\theta}$. To investigate this discrepancy, we plot the SC distribution for both of these ICs in Figure 2.3. Here, the IC seed area for A) IC 8 and B) IC 4 is shown in red, and its SC distribution to the rest of the brain is shown in blue. We observe that IC has a high degree of within-IC structural connectivity, while IC 4 has a low degree of within-IC structural connectivity relative to the rest of the brain. Permutation testing reveals the the strength of SC for these two ICs is significantly different ($p<0.0001$).

Next, we examine the difference in strength of SC for each IC between the control and MDD subject groups; the results are shown in Table 2.3. This table shows each group's mean $\hat{\theta}$ by IC, along with the bootstrap-based 95% confidence intervals and p-values (uncorrected) for the group difference in θ; p-values are calculated based on both the bootstrap SE and permutation testing. The mean $\hat{\theta}$ values are very similar for the control and MDD groups, so it is not surprising that none of the ICs show a significant between-group difference. This indicates that for the

Table 2.2: Results of hypothesis testing for controls

IC		Mean($\hat{\theta}_\ell$)	$\hat{SE}_B(\hat{\theta}_\ell)$	Bootstrap CI	p-value
2	motor	0.0071	0.0013	(0.0066, 0.0077)	<0.0001
3	FP	0.0081	0.0010	(0.0077, 0.0086)	<0.0001
4	**EC**	**0.0048**	0.0004	(0.0046, 0.0049)	<0.0001
5	FP	0.0077	0.0008	(0.0074, 0.0081)	<0.0001
8	**visual**	**0.0098**	0.0014	(0.0092, 0.0103)	<0.0001
10	EC	0.0052	0.0008	(0.0048, 0.0055)	<0.0001
11	visual	0.0085	0.0009	(0.0082, 0.0089)	<0.0001
12	motor	0.0058	0.0005	(0.0056, 0.0060)	<0.0001
13	DMN	0.0078	0.0016	(0.0071, 0.0085)	<0.0001

Figure 2.3: SC distribution for IC 8 (visual) vs IC 4 (exec control)



ICs that represent a normal resting state network, there is no substantial difference in the underlying structural connectivity between healthy controls and patients with MDD.

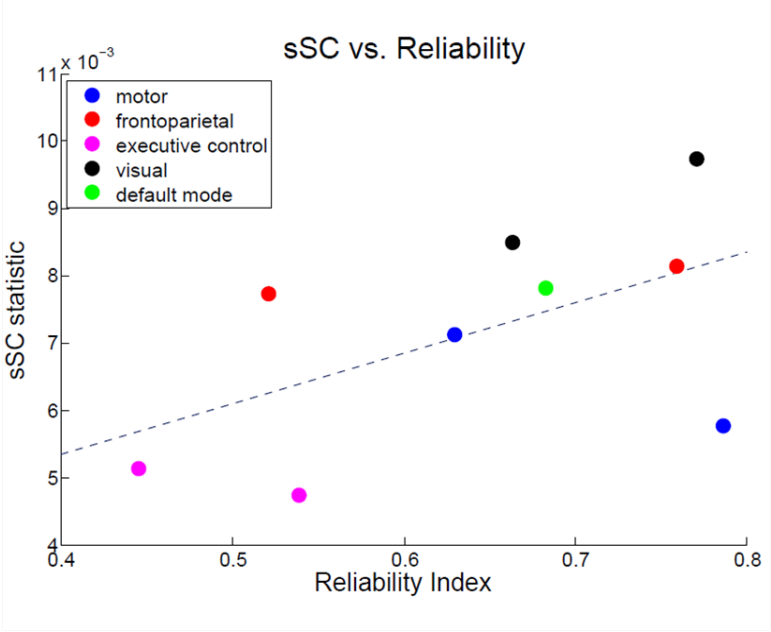
Finally, we investigated whether strength SC informs the reliability of the functional networks estimated by data-driven methods like ICA. We plot the strength of SC measure $\hat{\theta}_\ell$ vs our proposed reliability index R_ℓ for each IC ℓ, and find that these measures are positively associated (see Figure 2.4). That is, ICs with stronger underlying structural connectivity are more likely to be consistently estimated by

Table 2.3: Results of hypothesis testing for MDD vs. controls

IC		Controls $\hat{\theta}$ Mean	Controls $\hat{\theta}$ SE_b	MDD $\hat{\theta}$ Mean	MDD $\hat{\theta}$ SE_b	95% CI (MDD - con)	Bootstrap p-value	Perm. test p-value
2	motor	0.0071	0.0013	0.0070	0.0012	(-0.0009, 0.0005)	0.702	0.698
3	FP	0.0081	0.0010	0.0079	0.0012	(-0.0009, 0.0005)	0.525	0.562
4	EC	0.0048	0.0004	0.0047	0.0004	(-0.0003, 0.0002)	0.820	0.825
5	FP	0.0077	0.0008	0.0072	0.0010	(-0.0011, -0.0001)	0.018	0.054
8	visual	0.0098	0.0014	0.0095	0.0016	(-0.0012, 0.0007)	0.639	0.608
10	EC	0.0052	0.0008	0.0053	0.0008	(-0.0003, 0.0006)	0.451	0.475
11	visual	0.0085	0.0009	0.0080	0.0008	(-0.0010, 0.0000)	0.084	0.077
12	motor	0.0058	0.0005	0.0060	0.0007	(-0.0002, 0.0006)	0.277	0.299
13	DMN	0.0078	0.0016	0.0073	0.0012	(-0.0015, 0.0003)	0.251	0.252

the ICA algorithm. This suggests that we can leverage SC information from DTI data to inform the FC networks estimated from fMRI data.

Figure 2.4: Strength of SC is associated with IC reliability



2.6 Discussion

We integrated information from the fMRI and DTI data modalities to calculate a novel statistical measure of the strength of SC (sSC) underlying a functional network. Our simulation studies and data application demonstrated the utility of the sSC measure, and we found a positive association between sSC and the reliability of functional networks estimated by ICA.

Based on this finding, future work could include development of an SC-constrained ICA algorithm, which would leverage information from DTI data to estimate a more reliable and biologically plausible set of functional networks from fMRI data. We might also consider two modifications to the strength of SC measure. 1) Rather than standardizing the measure with respect to a maximum probability of 1, we could determine a more realistic maximum probability based on the amount of white matter in the IC. 2) Controlling for the average SC of a component is important, because different brain regions might have varying amounts of SC traffic, but this could unfairly penalize sSC in components with high underlying SC. Instead we might control for SC expected by chance, which could be determined by running tractography on a DTI image with a spherical tensor at each voxel. This way of measuring the "baseline" SC could account for connectivity observed by chance.

There has been an enormous amount of research devoted to FC over the past several years, with substantial interest focusing on resting-state FC. A set of resting state networks (RSNs) have been consistently identified in these investigations (Smith et al., 2009; Damoiseaux et al., 2006; Laird et al., 2011), most prominently the default mode network (DMN) (Buckner et al., 2008). While our work focused on resting-state connectivity, our proposed methods extend to other studies involving task-related fMRI.

Depression is a serious mental disorder affecting more than 20 million people

in the US and roughly 121 million people worldwide, according to the WHO. Due to the complexity of this disorder and its varied forms, its mechanisms are not fully understood. Functional and structural neuroimaging play critical roles in advancing our knowledge about major depression and other mental disorders. Our proposed methods stand to make a significant impact by improving our understanding of the neural representations of MDD, concentrating largely on the functional and structural relationships between different brain regions. Our research may have a long-term impact that is even more profound, since our proposed methods may generalize to studies of brain connectivity for other mental and neurological disorders, as well as to treatment studies.

Chapter 3

Topic 2: A joint model for functional and structural connectivity across the whole-brain network

3.1 Introduction

The network analysis approach has recently become a popular technique for modeling functional and structural connectivity across the whole brain. As of June 2016, a *Google Scholar* search of ”brain connectivity network analysis” yields 472,000 results. Adapted from graph theory, the network approach models the whole brain and it’s connections as a system of ”nodes” (brain regions) and ”edges” (pairwise connections between those regions) (Sporns et al., 2000). Network studies of psychopathology rely on the notion that dysfunctional nodes and edges result in abnormal signaling, which results in differences in brain connectivity networks. Since most major brain disorders involve dysfunction of processes that are distributed across multiple brain regions, rather than being constrained to one individual brain area, the network approach can provide important insights on global brain connectivity and organization in psychopathology (Menon, 2011).

A typical neuroimaging network analysis starts by parcellating the brain into regions, then quantifying a measure of association between all pairs of regions to produce a connectivity matrix (Sporns et al., 2000). For functional connectivity, the association between each region pair can be measured using pairwise Pearson’s correlations or partial correlations between each regions’ fMRI BOLD times series. For structural connectivity, the association between region pairs can be represented by the pairwise probability of SC or fiber counts, as estimated by a probabilistic or deterministic tractography procedure, respectively, using DTI data.

In this chapter, we present a novel multimodal approach that uses fMRI and DTI data to explore the relationship between functional and structural connectivity at each edge in the whole brain network. We use the expectation-maximization (EM) algorithm to estimate the model parameters by maximizing the data likelihood. It is our hypothesis that edges within a functional module will show a stronger association between SC and FC, relative to edges connecting regions from

different functional modules. We also explore whether the SC-FC relationship varies depending on how we represent FC (e.g. correlation vs. partial correlation). Since partial correlations are said to yield more accurate estimates of the true FC network (Smith et al., 2011), we might expect to see a stronger SC-FC association when FC is measured in this way. We will illustrate the application of our method to a resting-state fMRI and DTI dataset, and conduct simulation studies to evaluate the performance of our method.

3.2 Data

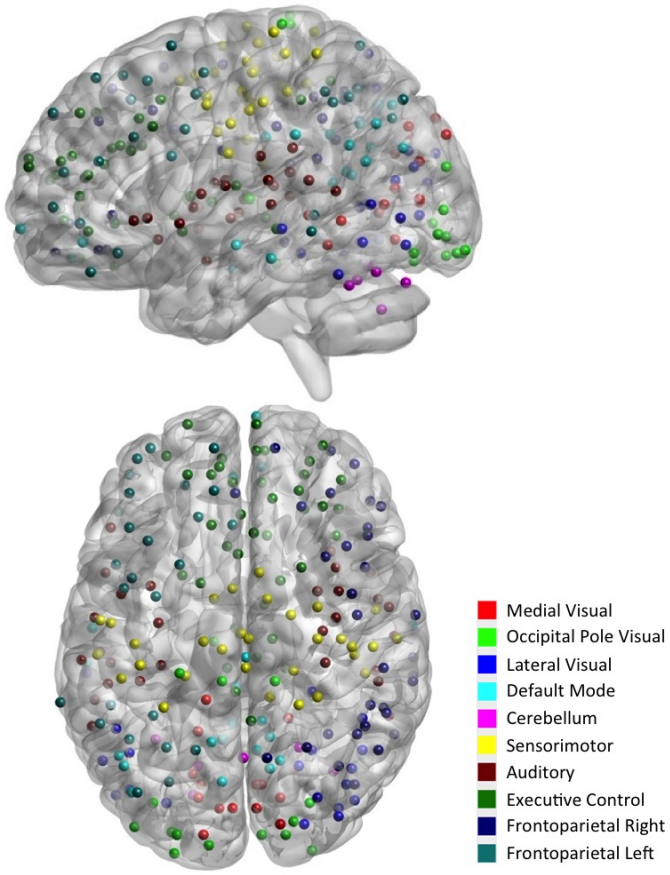
We use the dataset described in section 1.6, that contains resting-state fMRI (rs-fMRI) and DTI scans for 20 subjects with major depressive disorder (MDD) and 20 healthy controls.

3.2.1 Functional Connectivity (FC) Matrix Construction

One major challenge in brain network construction lies in defining the nodes of the network (Zalesky et al., 2010). While Stanley et al. (2013) recommends using individual voxels as network nodes, this results in an extremely high-dimensional connectivity matrix and can be biased by artificially strong local connections. Moreover, a voxel-based network can be highly variable across subjects due to the difficulty of aligning brain locations across the brains of different individuals. On the other hand, coarse anatomical atlases like the commonly used AAL atlas (Tzourio-Mazoyer et al., 2002) include large regions that are likely to contain multiple functional subregions, which can obscure the true properties of the network. For our network analysis, we adopt an intermediate whole-brain parcellation scheme defined by 264 putative functional regions (Power et al., 2011) that span the cerebral cortex, subcortical structures, and the cerebellum. This system of nodes was de-

termined using a meta-analysis of task-based fMRI studies combined with resting-state functional connectivity mapping techniques. In this parcellation, each node is a 5mm-radius sphere (81 voxels) in standard 2mm MNI space representing a putative functional region, and the collection of nodes provides reasonable coverage of the whole brain (see Figure 3.1). Relative to the voxel-level or large anatomical atlas parcellations, this system provides a good balance of spatial resolution and dimension reduction.

Figure 3.1: The 264-node system, organized by functional module.



Once we have chosen a parcellation scheme, we must next obtain a single representative time series for each region. Thus, for each individual subject and region, we perform a singular value decomposition (SVD) in the time domain to the $T \times V$ matrix of fMRI data, where T is the number of scans and V is the number of voxels

in the region (V=81). We extract the first right singular vector, which represents the dominant time series for that region. The resulting 264 temporal profiles summarize the neural activity at each node in our network system. To generate the 264 $\times$ 264 FC matrix, we measure the association between a pair of regions' BOLD fMRI time series in three ways: 1) using Pearson's correlations (henceforth referred to as "full correlation"), 2) using partial correlations at the 100% density level, and 3) using partial correlations at the 50% density level.

Full correlations are simple to calculate, but they only reflect the marginal association between network nodes, rather than the true or direct functional connection between them. We may observe a large correlation between a pair of nodes because of their common connections to a third-party node, even if the two nodes lack a direct connection. Partial correlations, on the other hand, measure the direct FC between two nodes, by estimating their correlation after regressing out effects from all the other nodes in the network, thus eliminating spurious effects (Smith et al., 2011). Through simulation studies, Smith et al. (2011) compared the performance of several FC network modeling methods, and found that partial correlation performed well and showed high sensitivity to detect true functional connections. Full correlations are still a reasonable and widely-used measure of FC in network models for fMRI data, however (Zalesky et al., 2012).

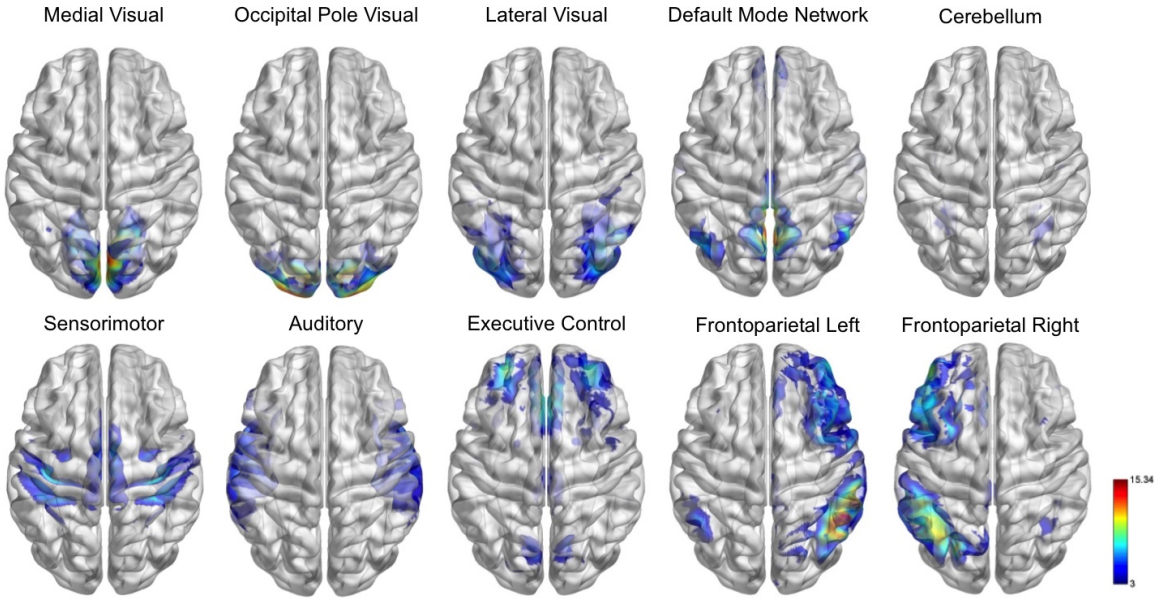
Despite its advantages in measuring direct FC, the application of partial correlation in the neuroimaging community has been limited, primarily because its calculation is computationally difficult, and requires choosing a tuning parameter to control the sparsity. Wang et al. (2016) presents an efficient and reliable approach to estimating the partial correlation matrix, and guides the choice of tuning parameter to achieve the desired matrix density level. We use the associated *DensParCorr* R package to estimate the partial correlation matrix at the 100% and 50% density levels (which corresponds to a tuning parameter λ of 0.001 and 0.025, respectively).

In order to understand the functional organization of the nodes in our network, we assigned them to one of ten functional modules, corresponding to the ten major resting state networks (RSNs) defined by Smith et al. (2009). The RSN maps (shown in Figure 3.2), determined by ICA decomposition of a large database of activation studies (BrainMap) and resting state fMRI data, are circuits whose BOLD activity is temporally coherent during both task activity and at rest. The functional modules include a medial visual network (*Med Vis*, 15 nodes), an occipital pole visual network (*OP Vis*, 15 nodes), a lateral visual network (*Lat Vis*, 19 nodes), the Default Mode network (*DMN*, 20 nodes), the cerebellum (*CB*, 6 nodes), a sensorimotor network (*SM*, 31 nodes), an auditory network (*Aud*, 29 nodes), an executive control network (*EC*, 39 nodes), and a right and left frontoparietal network (*FPR* and *FPL*, 32 and 26 nodes, respectively). To determine the module membership at each node, we identified the RSN z-statistic map with the largest value in the location of the node, above a chosen threshold ($z > 3$). Thirty two of the 264 nodes were not strongly associated with any RSN maps, and were therefore not included in our analysis. All brain visualizations were created using BrainNet Viewer (Xia et al., 2013).

3.2.2 Structural Connectivity (SC) Matrix Construction

The 264 nodes described in section 3.2.1 are centered in gray matter, so we needed to identify the nearby white matter fibers for each node in order to evaluate SC between region pairs. Rudie et al. (2013) performed a similar analysis with the 264-node system, and suggests the following procedure. Using the same set of 264 MNI coordinates from Power et al. (2011), we dilate each spheres to have a 10mm-radius to use for evaluating SC. Enlarging spheres to this size (515 voxels, compared to 81 voxels for functional spheres) ensures inclusion of white matter voxels, as defined by FSL's white matter tissue prior mask, thresholded at 50%.

Figure 3.2: Ten major resting state networks (RSNs)



On average, 68.4% of the voxels in each region are in white matter.

In order to evaluate the SC between each pair of regions, we perform region-to-region probabilistic tractography, implemented in FSL (Behrens et al., 2007). This procedure initiates several streams (5000 in our analysis) from each voxel in the seed region mask, and the streams each choose a path based on the principal direction of diffusion at each voxel, while avoiding passing through CSF. The proportion of streams that reach the target region represents the probability of SC for the seed-target region pair. We normalize this number by the *waytotal* count (i.e. the number of viable streams initiated from the seed region) to yield a more accurate measure of SC. The probabilistic nature of this algorithm can yield asymmetric region-to-region SC values. We impose symmetry in the final SC matrix by taking the average of the two directional measures for each region pair.

3.3 Methods

3.3.1 Joint model of SC and FC

We develop a hierarchical modeling framework to assess the link between SC and FC at each edge in the network.

3.3.1.1 Level 1

The first level models the observed FC and SC values and the variability due to repeated sessions/experiments. Consider R_{ikj}, the observed FC (as measured by full or partial correlation) at edge k, for subject i ($i = 1, ..., n$), and fMRI scanning session j ($j = 1, ..., J$). (If an fMRI data does not contain multiple scanning sessions per subject, one may consider dividing the time series of a single scan into equally-spaced intervals, to multiple pseudo-sessions. Since our fMRI dataset contains only one scanning session, we split the time course of the single scan into three even sections (J=3) of 50 time points each.) We apply the Fisher's-z transformation to transform R_{ikj} to the scale of all real numbers, and model the observed FC ($\tilde{R}_{ikj}$) as the sum of the subject-specific FC (f_{ik}) plus some normally distributed measurement error term (ε^F_{ikj}); i.e.

$$\text{arctanh}(R_{ikj}) = \tilde{R}_{ikj} = f_{ik} + \varepsilon^F_{ikj}$$

where

$$\varepsilon^F_{ikj} \sim N(0, \sigma^2_k)$$

In addition, consider Y_{ikl}, the observed probability of SC (as measured by a probabilistic tractography procedure) for edge k, subject i, and repeated tractography experiment l ($l = 1, ..., L$). (In our analysis, we repeat the probabilistic tractography experiment twice, under different random seed arguments, to get two

replications (L=2) per subject.) We apply the logit function to transform p_{ikl} to the scale of all real numbers, and model the observed SC (p'_{ikl}) as teh sum of the subject-specific SC (s_{ik}) plus some normally distributed measurement error term (ε^S_{ikl}); i.e.

$$\text{logit}(p_{ikl}) = p'_{ikl} = s_{ik} + \varepsilon^S_{ikl}$$

where

$$\varepsilon^S_{ikl} \sim N(0, \gamma^2_k)$$

3.3.1.2 Level 2

In the second level models the relationship between FC and SC using a linear link function, inspired by the function used to link longitudinal and time-to-event data presented in Tsiatis and Davidian (2004). Preliminary empirical evidence presented later in section 3.5 also supports the use of a linear model (see Figure 3.7). We propose to jointly model structural and functional connectivity for edge k in the network using the following link function:

$$f_{ik} = \lambda_k s_{ik} + \beta_k$$

Since FC is driven by SC, we model the subject's edge-level FC (f_{ik}) in terms of a linear model based on that subject's edge-level SC (s_{ik}). However, since FC is not always fully explained by SC, we include an intercept term to represent the baseline level of FC. Here, λ_k is the main parameter of interest, as it represents the FC-SC association at edge k, and β_k represents the baseline FC at edge k.

3.3.1.3 Level 3

Finally, in the third level, we model the subject-specific SC (s_{ik}) as the sum of the population SC (s_k) plus some normally distributed between-subject variabil-

ity term (e_{ik}); i.e.

$$s_{ik} = s_k + e_{ik}$$

where

$$e_{ik} \sim N(0, \tau_k^2)$$

Note that the current model framework requires replicated samples to perform the estimation. By requiring repeated experiments (e.g. multiple fMRI sessions and tractography replications), we can better account for the variability in SC and FC, and thus achieve more reliable estimates.

3.3.1.4 EM algorithm

We can use the EM algorithm (Dempster et al., 1977) and the maximum likelihood framework to estimate the set of parameters $\boldsymbol{\theta} = \{\lambda_k, \beta_k, \sigma_k^2, \tau_k^2, \gamma_k^2, s_k\}$, using the observed FC and SC data, $\boldsymbol{Y} = \{\tilde{R}_{ikj}, p'_{ikl}\}$, and latent variable s_{ik}.

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \log P(\boldsymbol{Y}, s_{ik}|\boldsymbol{\theta})$$

To facilitate the estimation of $\boldsymbol{\theta}$, we make the simplifying assumption of independence between pairwise connections, and estimate the model parameters separately for each edge k in the network. By combining all the elements of the model, we construct the complete log-likelihood function of all observed and hidden variables for each edge k:

$$\begin{aligned}
\ell(\boldsymbol{\theta}|\boldsymbol{Y}, s_{ik}) = & \\
&= \log(P(\boldsymbol{Y}|s_{ik}, \boldsymbol{\theta}) \times P(s_{ik}|\boldsymbol{\theta})) \\
&= \log\left(\prod_{i=1}^{n}\left[\prod_{j} f(\tilde{r}_{ikj}|f_{ik}, \sigma_k^2) \times \prod_{j'} f(p'_{ikl}|s_{ik}, \gamma_k^2) \times f(s_{ik}|s_k, \tau_k^2)\right]\right)
\end{aligned}$$

$$= \log\left(\prod_{i=1}^{n}\left[\prod_{j}\frac{1}{\sqrt{2\pi}\sigma_k}e^{\left(-\frac{[\tilde{r}_{ikj}-(\lambda_k s_{ik}+\beta_k)]^2}{2\sigma_k^2}\right)}\right.\right.$$
$$\left.\left.\times\prod_{l}\frac{1}{\sqrt{2\pi}\gamma_k}e^{\left(-\frac{(p'_{ikl}-s_{ik})^2}{2\gamma_k^2}\right)}\times\frac{1}{\sqrt{2\pi}\tau_k}e^{\left(-\frac{(s_{ik}-s_k)^2}{2\tau_k^2}\right)}\right]\right)$$
$$= -\left(\frac{J\lambda_k^2}{2\sigma_k^2}+\frac{L}{2\gamma_k^2}+\frac{1}{2\tau_k^2}\right)\sum_{i=1}^{n}s_{ik}^2$$
$$+\sum_{i=1}^{n}s_{ik}\left[\frac{1}{\sigma_k^2}\sum_{j=1}^{J}\lambda_k(\tilde{r}_{ik}-\beta_k)+\frac{1}{\gamma_k^2}\sum_{l=1}^{L}p'_{ikl}+\frac{s_k}{\tau_k^2}\right]+\text{const}$$

E-step:

The E-step function is the expected value of the log-likelihood function, with respect to the conditional distribution of $s_{ik}|\boldsymbol{Y}$. For iteration t, we fix the parameter estimates ($\boldsymbol{\theta^{(t)}}$) and update the explicit E-step function:

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)}) = E_{s_{ik}|\boldsymbol{Y},\boldsymbol{\theta}^{(t)}}[\ell(\boldsymbol{\theta}|\boldsymbol{Y},Z)]$$
$$= -\left(\frac{J\lambda_k^2}{2\sigma_k^2}+\frac{L}{2\gamma_k^2}+\frac{1}{2\tau_k^2}\right)\sum_{i=1}^{n}\overbrace{E(s_{ik}^2|\boldsymbol{Y},\boldsymbol{\theta}^{(t)})}^{E_{i2}}$$
$$+\sum_{i=1}^{n}\underbrace{E(s_{ik}|\boldsymbol{Y},\boldsymbol{\theta}^{(t)})}_{E_{i1}}\left[\frac{1}{\sigma_k^2}\sum_{j=1}^{J}\lambda_k(\tilde{r}_{ik}-\beta_k)+\frac{1}{\gamma_k^2}\sum_{l=1}^{L}p'_{ikl}+\frac{s_k}{\tau_k^2}\right]+\text{const}$$

where

$$\text{const} = -nJ\log(\sqrt{2\pi}\sigma_k)-nL\log(\sqrt{2\pi}\gamma_k)-n\log(\sqrt{2\pi}\tau_k)$$
$$-\frac{1}{2\sigma_k^2}\sum_{i=1}^{n}\sum_{j=1}^{J}(\tilde{r}_{ikj}^2+\beta_k^2-2\tilde{r}_{ikj}\beta_k)-\frac{1}{2\gamma_k^2}\sum_{i=1}^{n}\sum_{l=1}^{L}p'^2_{ikl}-\frac{ns_k^2}{2\tau_k^2}$$

and E_{i1}, E_{i2} are the conditional first and second moment of the latent variable s_{ik}

We show in Appendix B that

$$[s_{ik}|\boldsymbol{\theta}^{(t)}, \boldsymbol{Y}] \sim N(\mu^*_{ik}, \sigma^{*2}_{ik})$$

$$E_{i1} = E(s_{ik}|\boldsymbol{Y}, \boldsymbol{\theta}^{(t)}) = \mu^*_{ik}$$

$$E_{i2} = E(s^2_{ik}|\boldsymbol{Y}, \boldsymbol{\theta}^{(t)}) = [E(s_{ik}|\boldsymbol{Y}, \boldsymbol{\theta}^{(t)})]^2 + Var(s_{ik}|\boldsymbol{Y}, \boldsymbol{\theta}^{(t)}) = \mu^{*2}_{ik} + \sigma^{*2}_{ik}$$

where

$$\mu^*_{ik} = \frac{\left[\frac{\lambda_k(\tilde{r}_{ik\bullet} - J\beta_k)}{\sigma^2_k} + \frac{p'_{ik\bullet}}{\gamma^2_k} + \frac{s_k}{\tau^2_k}\right]}{\left[\frac{J\lambda^2_k}{\sigma^2_k} + \frac{L}{\gamma^2_k} + \frac{1}{\tau^2_k}\right]} \qquad \sigma^{*2}_{ik} = \frac{1}{\left[\frac{J\lambda^2_k}{\sigma^2_k} + \frac{L}{\gamma^2_k} + \frac{1}{\tau^2_k}\right]}$$

Notation: $\tilde{r}_{ik\bullet} = \sum_{j=1}^{J} \tilde{r}_{ikj}$ and $p'_{ik\bullet} = \sum_{l=1}^{L} p'_{ikl}$

M-step:

In the M-step of the algorithm, we fix the E-step function $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)})$ and update the parameter estimates ($\hat{\boldsymbol{\theta}}^{(t+1)}$).

$$\begin{aligned}
\hat{\boldsymbol{\theta}}^{(t+1)} &= \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\ Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)}) \\
&= \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \Bigg[-\left(\frac{J\lambda^2_k}{2\sigma^2_k} + \frac{L}{2\gamma^2_k} + \frac{1}{2\tau^2_k}\right) \sum_{i=1}^{n} E_{i2} \\
&\qquad + \sum_{i=1}^{n} E_{i1} \left(\frac{1}{\sigma^2_k} \sum_{j=1}^{J} \lambda_k(\tilde{r}_{ik} - \beta_k) + \frac{1}{\gamma^2_k} \sum_{l=1}^{L} p'_{ikl} + \frac{s_k}{\tau^2_k} \right) \\
&\qquad - nJ \log(\sqrt{2\pi}\sigma_k) - nL \log(\sqrt{2\pi}\gamma_k) - n \log(\sqrt{2\pi}\tau_k) \\
&\qquad - \frac{1}{2\sigma^2_k} \sum_{i=1}^{n} \sum_{j=1}^{J} (\tilde{r}^2_{ikj} + \beta^2_k - 2\tilde{r}_{ikj}\beta_k) - \frac{1}{2\gamma^2_k} \sum_{i=1}^{n} \sum_{l=1}^{L} p'^2_{ikl} - \frac{n s^2_k}{2\tau^2_k} \Bigg]
\end{aligned}$$

Explicit solutions for the parameter MLE equations are shown in Table 3.1.

Table 3.1: Maximum likelihood estimate equations based on EM algorithm

$$\hat{\lambda}_k^{(t+1)} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{J} E_{i1}(\tilde{r}_{ikj} - \beta_k^{(t)})}{J\sum_{i=1}^{n} E_{i2}}$$

$$\hat{\beta}_k^{(t+1)} = \frac{1}{nJ}\sum_{i=1}^{n}\sum_{j=1}^{J}(\tilde{r}_{ikj} - \lambda_k^{(t)} E_{i1})$$

$$\hat{\sigma_k^2}^{(t+1)} = \frac{1}{nJ}\sum_{i=1}^{n}\sum_{j=1}^{J}\left[E_{i2}\lambda_k^{(t)^2} - 2E_{i1}\lambda_k(\tilde{r}_{ikj} - \beta_k^{(t)}) + (\tilde{r}_{ikj} - \beta_k^{(t)})^2\right]$$

$$\hat{\tau_k^2}^{(t+1)} = \frac{1}{n}\sum_{i=1}^{n}\left[E_{i2} - 2s_k^{(t)}E_{i1} + s_k^{(t)^2}\right]$$

$$\hat{\gamma_k^2}^{(t+1)} = \frac{1}{nL}\sum_{i=1}^{n}\sum_{l=1}^{L}\left[E_{i2} - 2E_{i1}p'_{ikl} + p'^2_{ikl}\right]$$

$$\hat{s}_k^{(t+1)} = \frac{1}{n}\sum_{i=1}^{n} E_{i1}$$

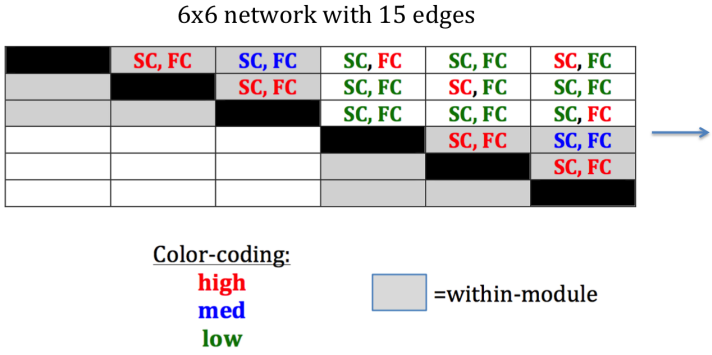
For each edge, indexed by k ($k = 1, ..., 34716$), we iterate between the E- and M-steps until convergence, which we define as $||\boldsymbol{\theta}^{(t+1)} - \boldsymbol{\theta}^{(t)}|| < 1 \times 10^{-8}$. For a network with 34716 unique edges, the total execution time ranges from 4-7 minutes, and the algorithm achieves a convergence rate of 99.99%.

3.4 Simulation Studies

We conduct simulation studies to evaluate the performance of the proposed joint model of FC and SC. Consider a simple brain network with 6 nodes and k=1,...,15 edges, organized in two functional modules (see Figure 3.3). We establish different levels of FC, SC, and FC-SC association depending on whether the edges are within- or between-module, based on the patterns we observe from real data. The

six within-module edges fall under two settings: high SC/high FC and medium SC/medium FC. We denote medium and high SC as having probability of connection around 0.3 and 0.6, respectively, and medium and high FC as having correlations around 0.5 and 0.7, respectively. Since we also expect these edges to show a strong SC/FC relationship, we set a high value of λ_k (0.25 or 0.5) at these connections.

Figure 3.3: Simulation setup for a 6x6 network.



	setting	λ_k
within-module edges	high SC, high FC	0.25
		0.25
		0.5
		0.5
	med SC, med FC	0.25
		0.5
between-module edges	high SC, low FC	0
		0.08
	low SC, high FC	0
		0.08
	low SC, low FC	0
		0
		0
		0.08
		0.08

The nine between-module edges fall under three settings: low SC/low FC, high SC/low FC, and low SC/high FC. We typically expect between-module nodes to show low connectivity, as empirical evidence suggests, although we occasionally observe high FC despite low SC, and vice versa. Here, we designate low SC and low FC to have values around 0.1 and 0.2, respectively. Since we also expect these between-module edges to show a weaker SC/FC relationship, we set a lower value of λ_k (0 or 0.08) at these connections. Guided by results from the data application, we choose values of σ_k^2=0.05, τ_k^2=0.2, and γ_k^2=0.01 for all edges $k = 1, ..., 15$. The remaining parameters, s_k and β_k, are assigned values such that each edge achieves its predetermined SC/FC setting. Note that for simplicity, we only consider *positive*

FC correlations.

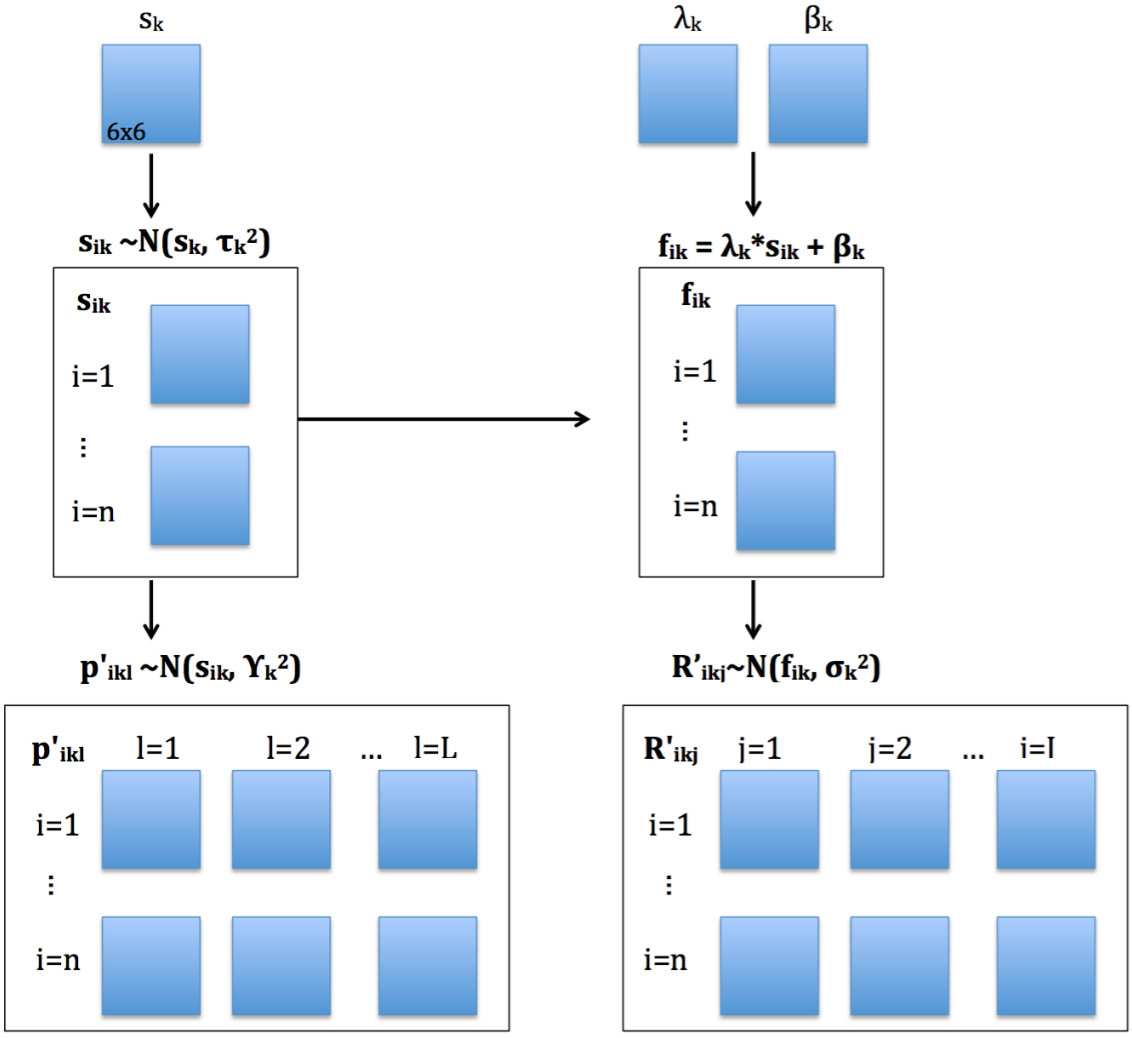
Based on the population-level parameter setting at each edge k, we can generate the subject-level means (s_{ik} and f_{ik}) and "observed data" ($\tilde{R}_{ikj}$ and p'_{ikl}). The process of generating simulated data is illustrated in Figure 3.4. For each simulation run, we will use the simulated data to estimate the edge-level parameters using the EM algorithm. We perform 1000 simulation runs under two sample size conditions (n=20 and n=50); for each run, 200 bootstrap samples of size n are drawn randomly with replacement, and used to generate an empirical distributions for the parameter estimates. As a result, we can estimate the bootstrap standard error (SE_{boot}) and coverage probabilities for 95% confidence intervals.

Tables 3.2 and 3.3 summarize the results based on 1000 simulation runs and two sample sizes, for the λ_k parameter at each the 15 edges, organized by SC/FC setting. (Tables B.1-B.4 in Appendix B show results for parameters β_k and s_k.) The ML estimates, bootstrap standard errors (SE_{boot}), and coverage probabilities are presented for each of the 15 edges in the network. Coverage probabilities are calculated from 95% confidence intervals in two ways: 1) based on the Wald-type interval using SE_{boot} and 2) using the bootstrap sample quantiles. (We do not present tables summarizing results for the three variance parameters, σ_k^2, τ_k^2, or γ_k^2, since they are not of interest.)

For each of the three parameters of interest, and under all configurations, we obtain average ML estimates with very low bias. For all parameters, SE_{boot} provides an accurate measure of variability, as compared with the empirical standard deviation of the mean ML estimates. Both measures of coverage probability are reasonably close to 95% coverage, although on average it is higher for the n=50 cases (93.6%) compared to the n=20 cases (92.1%).

The convergence rate for the EM algorithm was 100% under all configurations of sample size and edge settings. For n=20, the average computation time was

Figure 3.4: Generating simulated data.



fastest under the high SC conditions (45 iterations until convergence) compared to medium SC (117 iterations) and low SC (548 iterations), regardless of FC setting. The same pattern holds in the n=50 case, although the EM algorithm converges in about 10% fewer iterations.

Table 3.2: Simulation results for λ_k, n=20

setting	λ_k	$\hat{\lambda}_k$ mean	(SD)	SE_{boot}	CI cov prob (I)	CI cov prob (II)
high SC, high FC	0.25	0.246	(0.092)	0.088	93	92
	0.25	0.251	(0.090)	0.088	92	91
	0.5	0.503	(0.090)	0.091	93	92
	0.5	0.499	(0.089)	0.089	93	92
med SC, med FC	0.25	0.254	(0.089)	0.089	92	91
	0.5	0.500	(0.088)	0.091	93	92
high SC, low FC	0	0	(0.090)	0.087	91	91
	0.08	0.083	(0.093)	0.088	92	91
low SC, high FC	0	0.001	(0.084)	0.087	94	92
	0.08	0.08	(0.086)	0.087	94	93
low SC, low FC	0	-0.002	(0.088)	0.087	93	92
	0	0	(0.088)	0.087	92	92
	0	-0.006	(0.088)	0.085	93	91
	0.08	0.08	(0.087)	0.087	93	92
	0.08	0.081	(0.086)	0.087	92	91

3.5 Data Analysis

We apply the proposed joint model to a resting-state fMRI and DTI data from 20 healthy controls (see 1.6). FC and SC matrices were constructed using the 264 node system (Power et al., 2011) as described in Section 3.2.1. The 232×232 connectivity matrices, averaged across subjects and organized by functional module, are shown in Figure 3.5. SC is represented by a) the probability of SC between a pair of regions (estimated from probabilistic tractography), while FC is represented in three ways: b) full correlation, and partial correlation at c) 100% and d) 50% density. For the SC matrix, we observe that SC tends to be high for within-module edges, and low for between-module edges, as we might expect. Given their organization by functional module, it unsurprising that we observe a high degree of positive FC for for within-module edges, and primarily low or negative FC for between-module

Table 3.3: Simulation results for λ_k, n=50

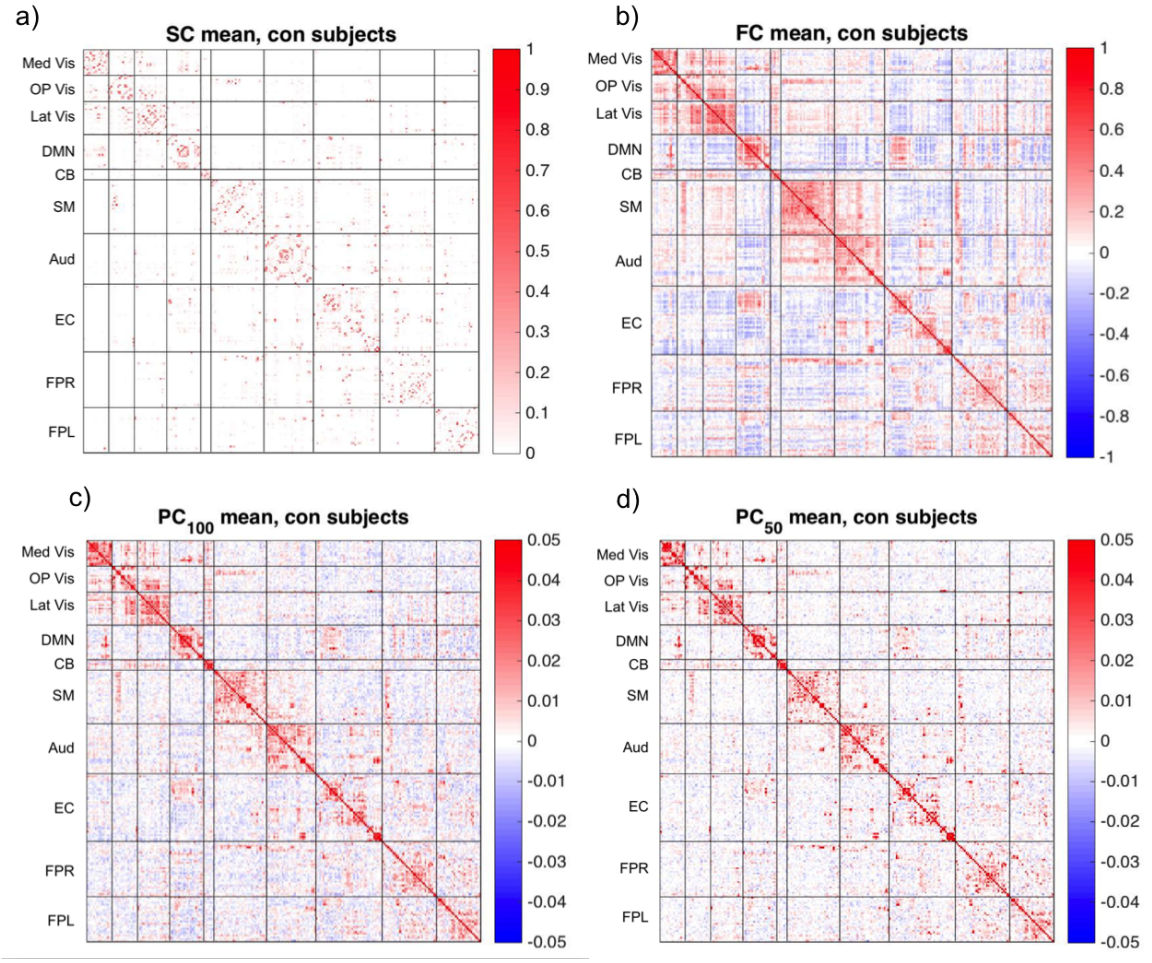
setting	λ_k	$\hat{\lambda}_k$ mean	(SD)	SE_{boot}	CI cov prob (I)	CI cov prob (II)
high SC, high FC	0.25	0.251	(0.052)	0.052	94	94
	0.25	0.251	(0.053)	0.052	92	93
	0.5	0.502	(0.054)	0.053	93	92
	0.5	0.498	(0.053)	0.053	95	95
med SC, med FC	0.25	0.249	(0.051)	0.052	95	94
	0.5	0.500	(0.054)	0.053	93	93
high SC, low FC	0	-0.001	(0.052)	0.052	95	94
	0.08	0.079	(0.052)	0.051	94	93
low SC, high FC	0	0.001	(0.053)	0.051	94	94
	0.08	0.079	(0.052)	0.051	94	94
low SC, low FC	0	-0.002	(0.051)	0.051	95	94
	0	0.002	(0.053)	0.051	93	93
	0	0.001	(0.053)	0.051	93	93
	0.08	0.081	(0.051)	0.052	94	93
	0.08	0.084	(0.051)	0.051	94	94

edges. It is also notable that compared to full correlations, partial correlations have a much smaller magnitude. In addition, the sparse partial correlation matrix (50% density) exhibits much fewer negative FC.

3.5.1 Measuring the edgewise FC-SC relationship

Although the group mean SC and FC values show a moderate linear relationship across the entire network of edges (see Figure 3.6), our model is instead intended to capture the relationship between SC and FC across subjects at the level of individual edges. As an example, in Figure 3.7 we provide the SC vs. FC plots for 3 edges in the network, displaying strong, moderate, and low FC-SC associations. Based on this empirical evidence, the linear link function appears to be a reasonable way to model the FC-SC relationship, although a larger sample size would

Figure 3.5: Mean connectivity matrices, for control subjects



give a clearer picture of the edge-level FC-SC association.

We apply our joint model to the fMRI and DTI, and estimate the ML estimates for the six model parameters using the EM algorithm. We also estimate the model parameters for 100 bootstrap resamples, to calculate the bootstrap standard error (SE_{boot}) for inference, which was shown in our simulations to be an adequate measure of variability. The λ_k parameter reflects the association between FC and SC at edge k, since it serves as the slope term in our proposed link function. Figure 3.8 displays the heatmaps of unstandardized positive $\hat{\lambda}_k$ values at each edge in the network, using a) full correlation, b) partial correlation at 100% density, and

Figure 3.6: SC vs FC, population-level associations

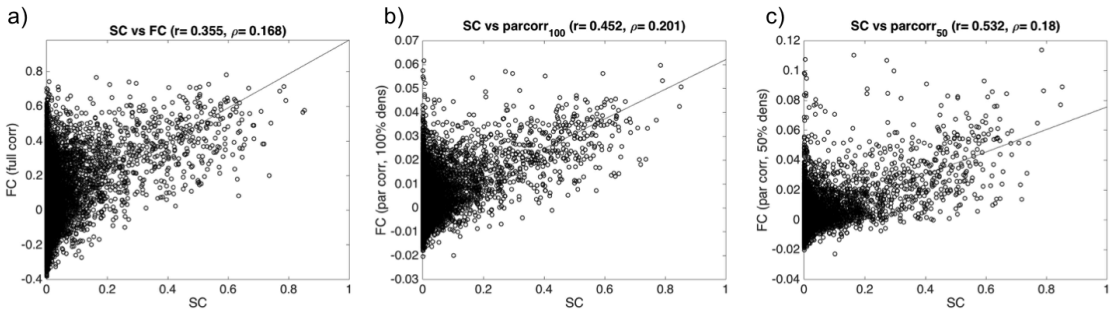
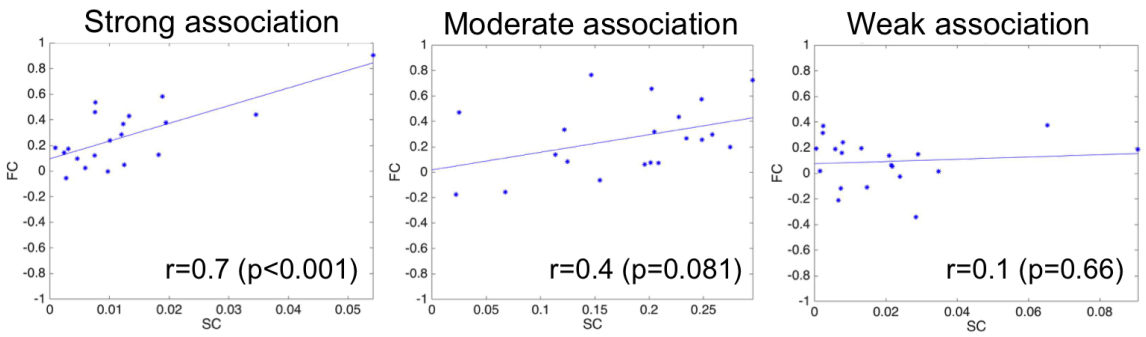


Figure 3.7: SC vs FC, edge-level associations



c) partial correlation at 50% density as measures of FC. (For interpretability, we only present the positive λ values). In each case, we see a block diagonal pattern, indicating that the FC-SC association is stronger for within-module edges, relative to between-module edges. This pattern is most pronounced in plot c), which uses sparse partial correlations as an FC measure. This provides some evidence that direct FC (as measured by partial correlations), has a stronger association with SC, relative to marginal FC (as measured by full correlations). However, when we use SE_{boot} to standardize and threshold to find the edges with significant $\hat{\lambda}_k$ values (uncorrected for multiple comparisons), the block-diagonal pattern is not as evident (see Figure B.1 in Appendix B). Furthermore, the magnitude of the estimated association measure depends heavily on the marginal distribution of FC and SC, as

seen by the difference in magnitudes of the $\hat{\lambda}_k$ plots in Figure 3.8. Thus, the results from this method are not directly comparable if we use full or partial correlation to measure FC.

Figure 3.8: Unstandardized $\hat{\lambda}_k$ results, using different measures of FC

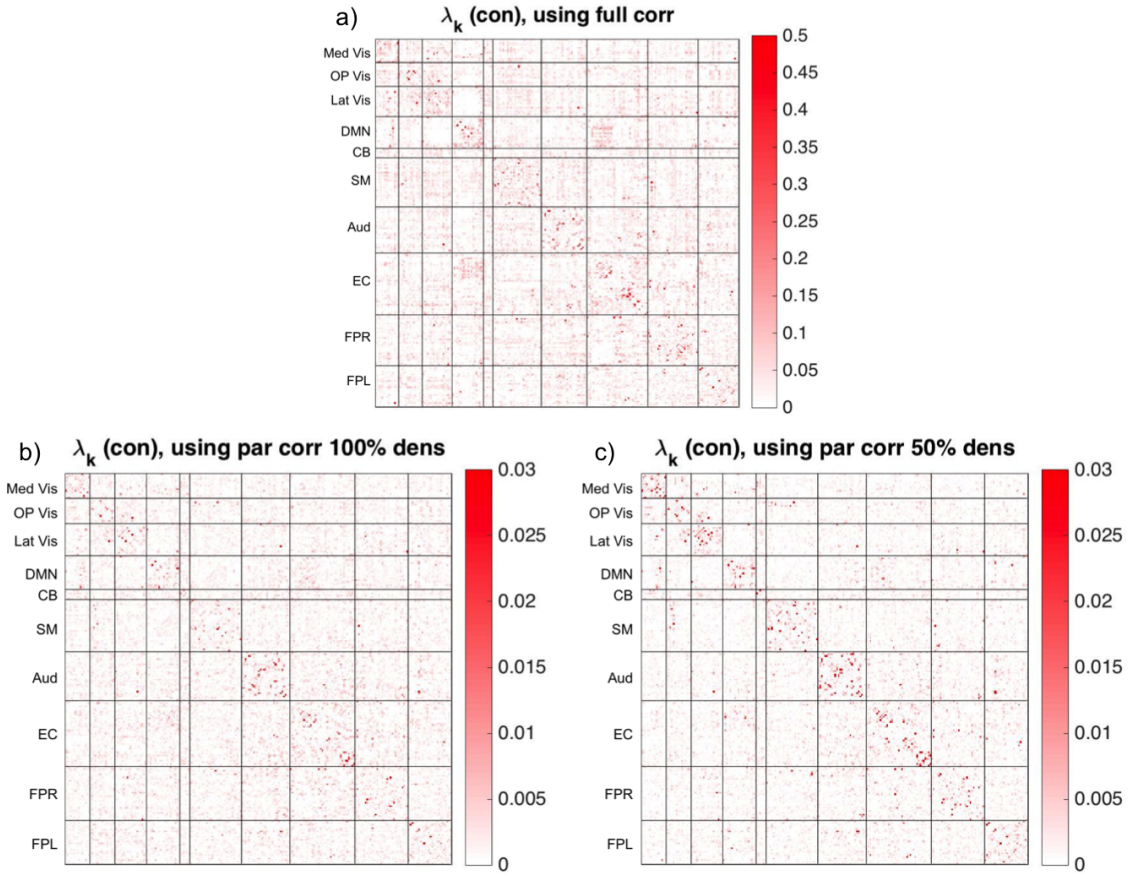


Figure 3.9 shows the edges with the a) 100 highest and b) 100 lowest positive $\hat{\lambda}_k$ values, when FC is measured using full correlation. We observe that the high positive $\hat{\lambda}_k$ values tend to be short-distance and located in within-module edges, while the low $\hat{\lambda}_k$ values tend to be long-distance and located in between-module edges. Furthermore, we observe that the 100 edges with high $\hat{\lambda}_k$ (i.e. high FC-SC association) tend to have both high underlying FC and SC, while edges with low $\hat{\lambda}_k$ (i.e. almost zero FC-SC association) tend to have underlying FC and SC

values near zero. We observe similar trends in $\hat{\lambda}_k$ when FC is measured by partial correlation.

Figure 3.9: 100 edges with the a) highest vs. b) lowest positive $\hat{\lambda}_k$ values

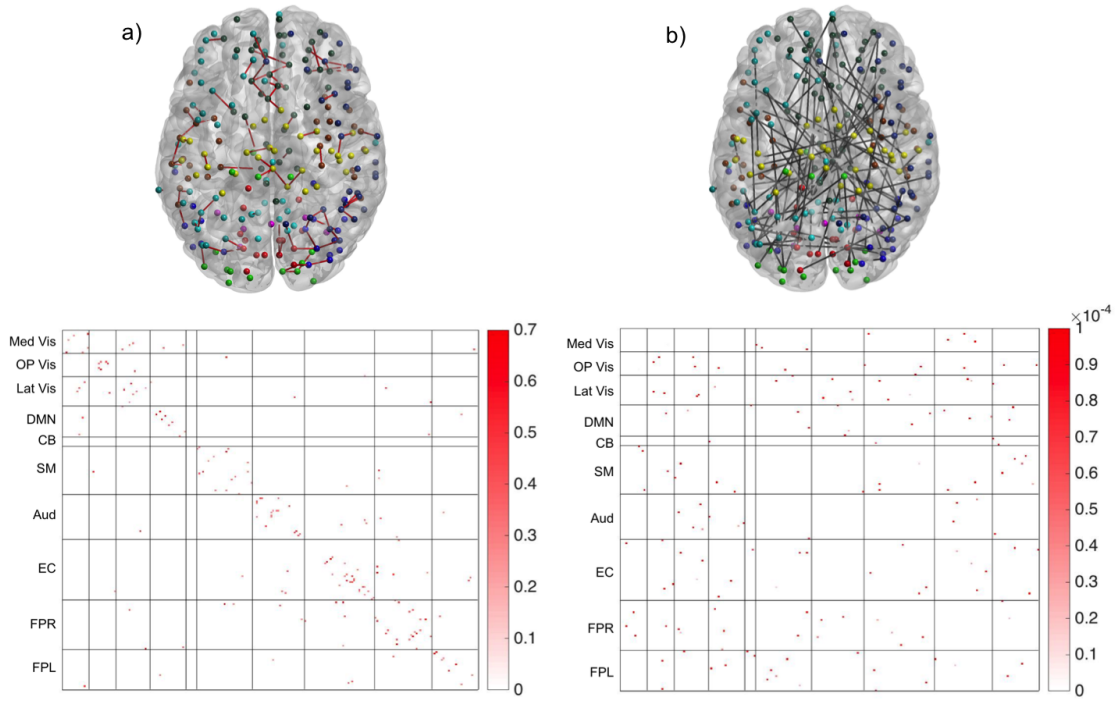
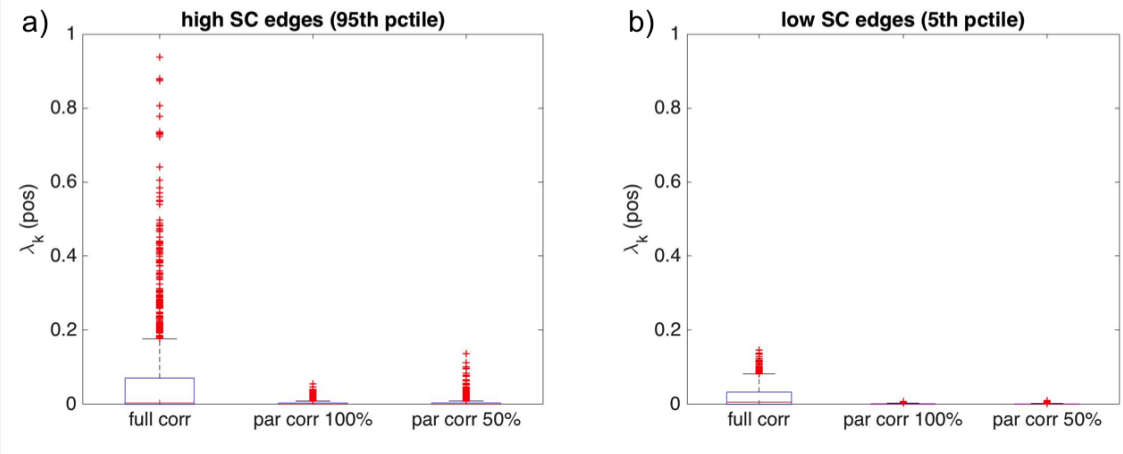


Figure 3.10 shows the distribution of positive $\hat{\lambda}_k$ values for edges with a) very high SC (95th percentile) vs. b) very low SC (5th percentile). Estimation is performed using three different measures of FC: full correlation, partial correlation at 100% density, and partial correlation at 50% density. We see that within each FC measure, the FC-SC association (captured by $\hat{\lambda}_k$) is much lower for edges with high SC compared with low SC, as expected. Comparing results from full vs. partial correlations, we see that the magnitude of $\hat{\lambda}_k$ is higher for the former, since full correlations tend to be higher than partial correlations (see Figure 3.5)

Figure 3.10: $\hat{\lambda}_k$ at edges with very high vs low SC values



3.5.2 Group comparison

Figure B.2 in Appendix B shows the mean SC and FC matrices for the 20 subjects with MDD. Comparing to the analogous matrices for healthy controls in Figure 3.5, we can see that the connectivity matrices for each subject group appear to be similar.

We perform permutation testing to compare the edge-level FC-SC association (using full correlation for FC) between controls and MDD patients; after controlling for multiple comparisons, we did not find any significant differences in λ_k between the two groups.

3.6 Discussion

Our proposed method offers an approach for modeling the FC-SC relationship across the whole-brain network. We employ the EM algorithm to obtain the maximum likelihood estimates of the model parameters, and use the bootstrap method to estimate their variability for further inference. Simulation studies under a variety of SC/FC settings demonstrate that our method yields accurate estimates of

the model parameters, and performs most efficiently for edges with high SC.

The application of our method to a resting-state fMRI and DTI dataset of 20 healthy controls further demonstrates the ability of our method to explain the relationship between SC and FC at each edge in the network. As hypothesized, the edges with the highest positive FC-SC association tend to be within-module and short-distance, while the edges with the lowest FC-SC association tend to be between-module and long-distance.

We also examine the FC-SC association when FC is measured via full correlation vs. partial correlation (at the 100% and 50% density levels). When using these different measures, we find that the results follow a consistent pattern in which the FC-SC association is much higher for within-module edges compared to between-module edges.

One strength of the model framework is that it requires replicated samples to perform the estimation. By requiring repeated experiments (e.g. multiple fMRI sessions and tractography replications), we can better account for the variability in SC and FC, and thus achieve more reliable model estimates.

3.6.1 Limitations and Future Considerations

Although we observe consistent patterns in the FC-SC association using different measures of FC (full vs. partial correlations), the scale of the association measure is heavily influenced by the marginal distribution of the connectivity measures. Thus, the magnitudes of $\hat{\lambda}_k$ are much higher when FC is measured by full correlations vs. when FC is measured by partial correlations. Unfortunately, this means that the association measure is not directly comparable when the method is applied using different connectivity measures. In Topic 3 (Chapter 4), we pursue a more robust measure of association that does not depend on the marginal distributions of the univariate connectivity values.

A limitation of our joint model is that it implies a linear relationship between FC and SC at each edge. A more flexible link function, such as that provided by a copula model may improve the joint model by allowing for a more flexible association between the modalities. This is the approach we pursue in the next chapter.

In addition, our method is limited in that it assumes independence between all edges, and estimates the model parameters at each edge separately. This assumption was made for feasibility of computation, but future models could incorporate a dependence structure between edges. By borrowing information from neighboring edges, we might improve estimation and yield a more realistic model of the function-structure association across the whole network.

Although our analysis did not uncover a significant difference between the control and MDD population, this is not unexpected, since our analysis employed a large number of edge-level tests (34716) and a small sample size (n=20 for each group). Eventually, we plan to validate our findings by using a larger dataset, to achieve higher statistical power.

Chapter 4

Topic 3: Using copulas to model the structure-function relationship in the brain

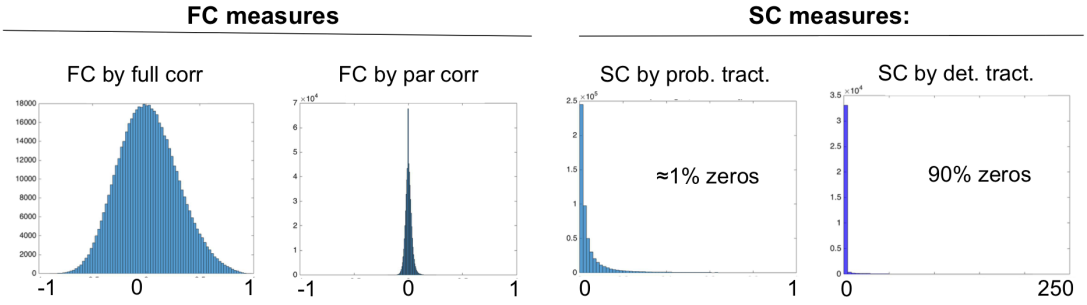
4.1 Introduction

In Topic 2, we presented a method of measuring the association of FC and SC that assumes a linear relationship. However, the classical method of linear association may not be appropriate for modeling the complex dependence structure between these random variables. Measuring dependence linearly with Pearson's correlation is known to have several pitfalls, and while it measures the overall strength of the association, it lacks information about how this association varies across the distribution. In addition, Pearson's correlation is strongly affected by extreme endpoints, and is not invariant under non-linear monotone increasing transformations of random variables (Embrechts et al., 1999). Non-parametric rank correlation coefficients like Kendall's tau (τ) and Spearman's rho (ρ) offer alternative measures of dependence that are invariant under monotonic transformations (Schweizer and Wolff, 1981), but they still only characterize the association with a single value, rather than describing the joint distribution of the random variables.

In addition, the scale of the association measure proposed in Topic 2 is heavily affected by the marginal distribution of the univariate connectivity values. For example, pairwise full correlations and partial correlations, both measures of FC, tend to have drastically different magnitudes. Likewise, SC can be measured using probabilistic or deterministic tractography, yielding pairwise probability values and fiber counts, respectively. These also have different distributions, with the fiber counts usually having a large proportion of absolute zero values. Figure 4.1 shows example histograms of these various connectivity measures. We aim to develop a more robust measure of the FC-SC association, that does not depend on the marginal behavior of the FC and SC measures.

For Topic 3, we consider the commonly overlooked theory of copulas (Joe, 2015; Nelsen, 2006) as a more flexible and robust approach to modeling the association between FC and SC at the joint probability density (pdf) level. Copulas are

Figure 4.1: Histograms different FC and SC measures



functions that link multivariate distribution functions to their univariate marginal distributions; this allows for simultaneous modeling of the dependence structure between random variables and separately, the arbitrary specification of each variable's univariate marginal distributions. Advantages of using copulas include 1) the ability to flexibly model both linear and non-linear dependence, 2) separate specification of marginal distributions, and 3) the capacity for modeling extreme endpoints. Implementation of the copula approach to measure the FC-SC association will be illustrated by application to an fMRI and DTI dataset, and via simulation studies.

4.1.1 Copulas

Copulas are a class of functions that "couple" multivariate distribution functions to their univariate marginal distribution functions, where the marginals are uniform on the interval [0, 1]. The copula completely characterizes the dependence between random variables, and allows the parameters of each marginal distribution function to be modeled separately. The main purpose of a copula is to disentangle the dependence structure between variables from their marginal distributions, so that the resulting association measure is not dependent on the scale of the marginals. The term "copula," meaning link or bond, was first used by Sklar (1959) to refer to

this class of distributions.

Historically, copulas have primarily been applied in the fields of finance (Cherubini et al., 2013) and survival analysis (Shih and Louis, 1995), although they are a versatile tool that can be applied in any situation involving two or more random variables. Recent studies have used copulas to examine associations in environmental health (Kostova et al., 2012) and climate science (Schölzel and Friederichs, 2008), and to make predictions in medical studies (Kumar and Shoukri, 2007, 2008). In the field of neuroscience, copulas have been used to describe the dependence structure between spike counts across a population of neurons (Onken et al., 2009; Jenison and Reale, 2004). In addition, Silva et al. (2014) describes the use of copulas to simulate multimodal neuroimaging data, which allows the user full control of the single-modality marginal distributions, as well as the type and level of association between modalities.

Consider two random variables X and Y, which represent structural and functional connectivity in the context of our model (see section 4.3). We denote the joint cumulative distribution function (CDF) as $H(x, y)$, and the marginal CDFs of X and Y as $F_X(x)$ and $F_Y(y)$, respectively. It is a well known result that the CDF of a random variable is uniformly distributed on the interval [0,1]. In order to use copulas, the random variables are transformed to uniformly distributed variables U and V via the marginal CDFs:

$$U = F_X(x) \sim U(0, 1)$$

$$V = F_Y(y) \sim U(0, 1)$$

There exists a bivariate copula function, C, such that

$$H(x, y) = C(F_X(x), F_Y(y)) = C(u, v)$$

The virtue of copulas stems from Sklar's theorem, which states that all continuously differentiable joint CDFs can be written as a multivariable function of their marginals (Sklar, 1959). Under the continuously differentiable assumption for $C(u, v)$ and $H(x, y)$, it can be shown that the joint probability density function (PDF) of X and Y can be written as

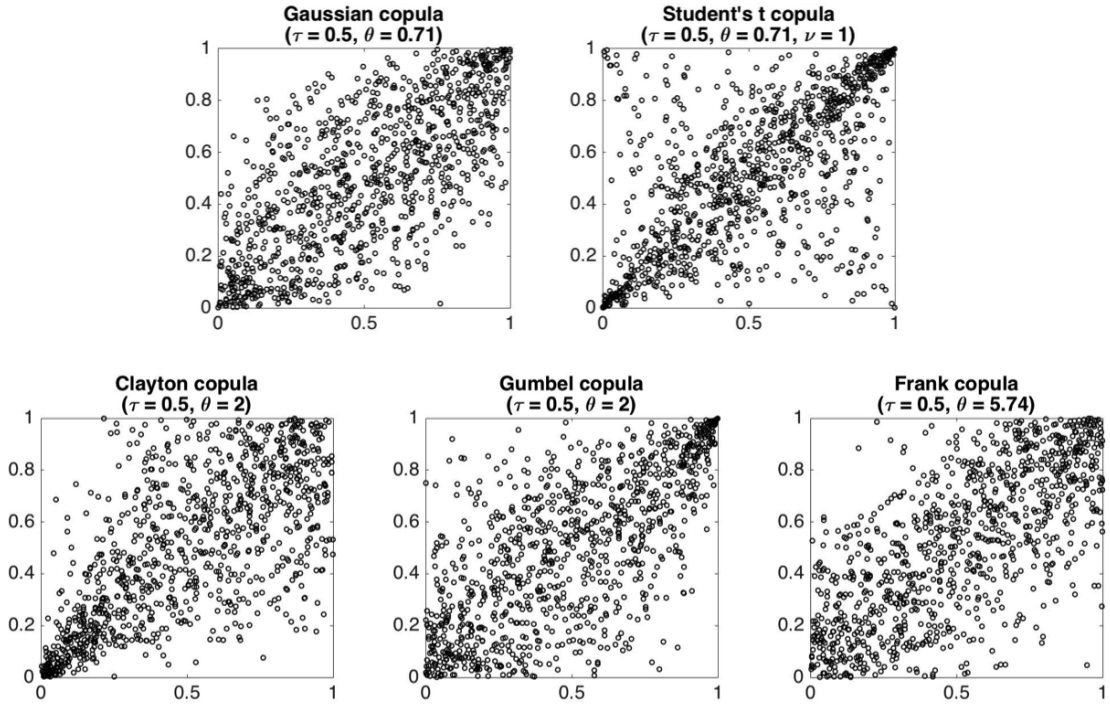
$$\begin{aligned} h(x, y) &= f_X(x) \times f_Y(y) \times c(u, v) \\ &= f_X(x) \times f_Y(y) \times c(F_X(x), F_Y(y)) \end{aligned}$$

Where $c(u, v)$ is the copula (joint) PDF of U and V. Therefore, the joint pdf $h(x, y)$ can be re-expressed as the product of the marginal pdfs times the copula pdf, and the dependence structure between X and Y is entirely captured by $c(u, v)$. This is the main advantage of using copulas, since the univariate marginal behavior of random variables can be modeled separately from their dependence.

There are many different types of copulas, which capture different types of dependence. The strength of dependence is determined by the corresponding parameter, θ, which has a one-to-one mapping to Kendall's τ. According to Silva et al. (2014), "the literature on the expected type of dependence between multimodal [neuroimaging] sources is virtually inexistent." In the following section, we outline 5 well-known copula functions: the Gaussian, t, Clayton, Gumbel, and Frank copulas (Joe, 2015; Nelsen, 2006), which capture a wide range of dependence structures and are easily implemented with the set of copula functions in Matlab (described here: http://www.mathworks.com/help/stats/copula-distributions-and-correlated-samples.html). Figure 4.2 shows the different dependence structures for data simulated under these five different copula functions. The plots in this figure are comparable because they each correspond to a Kendall's tau (τ) of 0.5. We will explore the use of these 5 copulas to investigate the association between FC and

SC.

Figure 4.2: Data simulated under different dependence structures (copulas), but with same Kendall's rank correlation ($\tau = 0.5$).



4.1.1.1 Elliptical copulas

Elliptical copulas, like the Gaussian and t copula are a class of symmetric copulas, so-called because the horizontal cross-sections of their bivariate pdfs take the shape of ellipses. Both Gaussian and t-copulas are easily parameterized by the linear correlation matrix, but only t-copulas yield dependence structures with tail dependence.

The **bivariate Gaussian copula** function is given by:

$$C(u, v; \theta) = \Phi_2(\Phi^{-1}(u), \Phi^{-1}(v); \theta)$$

where Φ^{-1} is the inverse CDF of the standard normal (i.e. probit function) and Φ_2 is the bivariate standard normal distribution with correlation coefficient θ.
See Figure C.1 in C to visualize the dependence structure of the Gaussian copula under different levels of θ.

The **bivariate t copula** function is given by:

$$C(u, v; \theta, \nu) = T_{2,\nu}(T_\nu^{-1}(u), T_\nu^{-1}(v); \theta)$$

where T_ν^{-1} is the inverse CDF of the t-distribution with ν degrees of freedom, and $T_{2,\nu}$ is the bivariate t-distribution with ν degrees of freedom and correlation coefficient θ. Not surprisingly, as the degrees of freedom parameter ν increases, the t copula approaches the corresponding Gaussian copula. See Figure C.2 in Appendix C to visualize the dependence structure of the t copula under different levels of θ.

For both the Gaussian and t copulas, the the linear correlation coefficient θ has the following one-to-one relationships with Kendall's tau (τ) and Spearman's rho (ρ):

$$\theta = \sin(\tfrac{\pi}{2}\tau) \quad \text{or} \quad \tau = \tfrac{2}{\pi}\text{asin}(\theta)$$

and

$$\theta = 2\sin(\tfrac{\pi}{6}\rho) \quad \text{or} \quad \rho = \tfrac{6}{\pi}\text{asin}(\tfrac{\theta}{2})$$

4.1.1.2 Archimedian copulas

The elliptical copulas described in section 4.1.1.1 are appropriate for modeling the dependence structure in symmetric data. Archimedian copulas, such as the Clay-

ton, Gumbel, and Frank copulas, have the form

$$C_\psi(u, v) = \psi(\psi^{-1}(u), \psi^{-1}(v))$$

where $\psi(t)$ is a generator function. For this class of copulas, the parameter θ describes the strength of dependence, and has a one-to-one mapping with Kendall's tau (τ).

Table 4.1 displays the generator function $\psi(t)$, copula function, Kendall's τ's relationship with θ, and range of possible θ values for each of the bivariate Clayton, Gumbel, and Frank copulas. Figures C.3, C.4, and C.5 in Appendix C show the dependence structure of the Clayton, Gumbel, and Frank copulas, respectively, under different levels of θ. As you can see from these plots, the Clayton copula allows for strong left tail dependence, while the Gumbel copula allows for strong right tail dependence between variables.

Table 4.1: Bivariate Archimedian Copulas

Copula	Generator fn $\psi(t)$	Copula function $C(u, v; \theta)$	Kendall's τ	Range of θ
Clayton	$t^{-\theta} - 1$	$(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$	$\frac{\theta}{\theta+2}$	$\theta > 0$
Gumbel	$(-\ln t)^\theta$	$\exp -[(-\ln u)^\theta + (-\ln v)^\theta]^{1/\theta}$	$1 - \theta^{-1}$	$\theta > 1$
Frank	$-\ln \frac{e^{-\theta t}-1}{e^\theta-1}$	$-\frac{1}{\theta}\ln(1 + \frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{(e^{-\theta}-1)})$	$1 - \frac{4}{\theta}[1 - D_1(\theta)]$	$\theta \in (-\infty, \infty)$

4.2 Data

We continue to use the rs-fMRI and DTI dataset described in section 1.6. As in Topic 2, we use the 264-node system (Power et al., 2011) (as described in sections 3.2.1 and 3.2.2) to define our FC and SC networks. For this analysis, we only consider full correlation as a measure of FC and use region-to-region probabilistic tractography to measure SC.

4.3 Methods

Copula functions allow the researcher to specify the dependence structure between random variables, and separately, to model their marginal distributions. As an extension to the methods discussed in Topic 2 (Chapter 3), we propose to use copulas to study the association between FC and SC in two ways. First, to examine whether the FC-SC association differs between edges located within-module vs. between-module, and second, to estimate the association between FC and SC at each edge in the whole-brain network.

4.3.1 Marginal distribution specification for SC and FC data

An important advantage of the copula is that it allows for any choice of marginal distribution for the random variables, X and Y. The copula is constructed on the assumption that the marginal distributions are known or estimated from the data.

The joint SC and FC model framework described by Venkataraman et al. (2012) proposes to model the edge-level SC as having two latent states (present or absent) and the edge-level FC as having three latent states (negative, positive, or no FC). Modeling SC as binary fails to capture the potential spectrum of SC strength values, however. Instead, based on empirical evidence, we suggest modeling the SC data as a mixture of four Gaussian components, representing 1) near-zero SC, 2) very low SC, 3) moderate SC, and 4) high SC. As prescribed by Venkataraman et al. (2012), we will model the FC data as a mixture of three Gaussian components, representing 1) negative FC, 2) near-zero FC, and 3) positive FC. To maintain the notation presented in section 4.1.1, we will denote the SC and FC random variables by X and Y, respectively.

Prior to estimating the model, we transform the SC and FC data. While a typical transformation of data in the range of [0, 1] is the logit or probit function, the SC

data measured by region-to-region probabilistic tractography may contain values of 0 and 1 that are meaningful. Instead, we use the arcsine transformation:

$$X = \arcsin(\sqrt{p})$$

where p is the observed edgewise probability of SC extracted from a probabilistic tractography procedure, and X is the transformed SC data. This function produces valid transformations for 0 and 1 values, and maintains the original shape of the data. For FC data, we employ the commonly-used Fisher's-z transformation:

$$Y = \text{arctanh}(r)$$

where r is the observed edgewise full correlation between two region's BOLD fMRI time series, and Y is the transformed FC data.

The marginal PDF of the transformed SC data (X) can be modeled as a mixture of 4 Gaussian components:

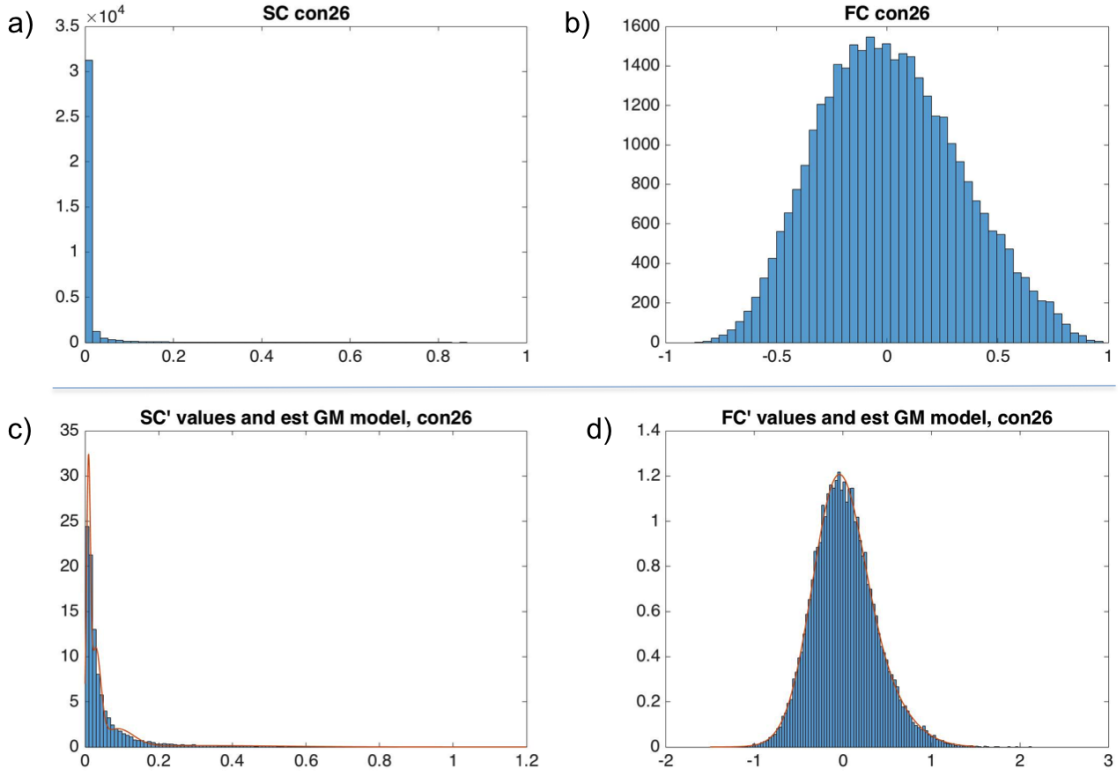
$$f_X(x) = \sum_{j=1}^{4} p_j \times N(x; \mu_j, \sigma_j^2) \quad \text{where} \quad \sum_{j=1}^{4} p_j = 1$$

Moreover,the marginal PDF of the transformed FC data (Y) can be modeled as a mixture of 3 Gaussian components:

$$f_Y(y) = \sum_{l=1}^{3} \pi_l \times N(y; \lambda_l, \tau_l^2) \quad \text{where} \quad \sum_{l=1}^{3} \tau_l = 1$$

To justify using a Gaussian mixture (GM) model for the SC and FC data, we preliminarily examine the distributions of X and Y using one control subject's SC and FC data. Plots a) and b) of Figure 4.3, show the distribution of the raw SC and FC data, respectively, across all edges in the network. (This data is taken from one randomly chosen control subject in Depression dataset.) We see that the SC data is heavily right skewed, with several values at or near zero. We also observe a

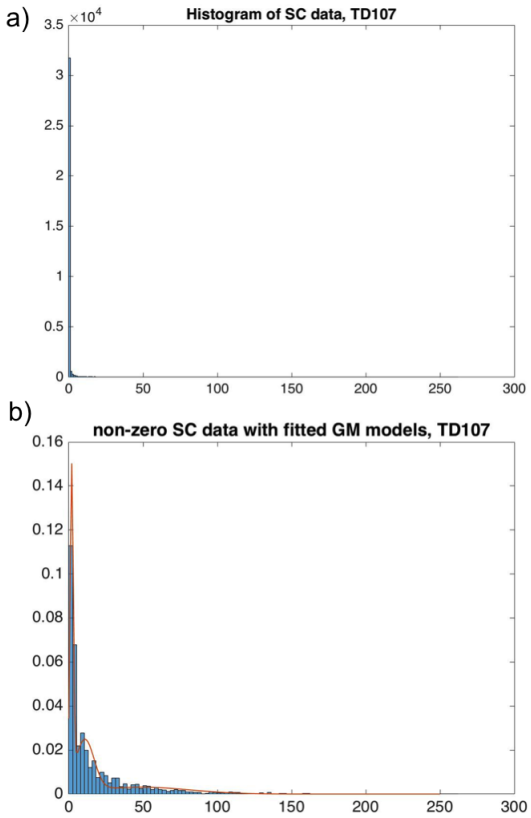
Figure 4.3: Histograms of SC and FC data for one control subject (con26) from Depression dataset



fairly normal distribution of FC values, centered around 0. Plots c) and d) show the transformed SC' and FC' data (X and Y), overlaid with the pdfs of the fitted GM models (in orange), respectively, demonstrating reasonably good fit to the data.

This method can also be adapted to cases where SC is measured deterministically, (section 1.4.3), in which region-to-region SC is measured using fiber counts rather than probabilities. In this case, the distribution of SC values will be similar, with a strong right skew, but there will be a higher spike at zero due to the large number of absolute zeros in the data. For example, the histogram in plot a) of Figure 4.4 shows a typical distribution of SC values as measured by fiber counts. (This data is derived from a randomly chosen typically developing (TD) subject in an rs-fMRI and DTI study of Autism (Brown et al., 2012).) In this dataset, about

Figure 4.4: Histogram of SC (fiber counts), for one TD subject (TD107) from Autism dataset



90% of edges have SC values of zero, so the proposed 4-component Gaussian mixture model fails to fit the data. When modeling SC data measured by fiber counts, we suggest substituting the first Gaussian component with the Dirac delta function $\delta(x)$ to represent the spike at zero. The remaining non-zero SC data can still be treated as a mixture of 3 Gaussian components, as shown in plot b) of Figure 4.4.

The Gaussian mixture (GM) models can be estimated from the SC and FC data using a set of Matlab functions (see: http://www.mathworks.com/help/stats/gaussian-mixture-distribution-1.html). Once these marginal PDF parameters have been estimated from the data, we must next obtain the marginal CDFs to use as uniformly distributed variables U and V in the copula function. Variables U and

V are derived as follows, and can be easily calculated using the *normcdf* function in Matlab.

$$
\begin{aligned}
U &= F_X(x) \\
&= \int_{-\infty}^{x} \sum_{j=1}^{4} p_j \times N(t; \mu_j, \sigma_j^2) dt \\
&= p_1 \int_{-\infty}^{x} N(t; \mu_1, \sigma_1^2) dt + p_2 \int_{-\infty}^{x} N(t; \mu_2, \sigma_2^2) dt \\
&\qquad + p_3 \int_{-\infty}^{x} N(t; \mu_3, \sigma_3^2) dt + p_4 \int_{-\infty}^{x} N(t; \mu_4, \sigma_4^2) dt \\
&\sim U(0,1)
\end{aligned}
$$

$$
\begin{aligned}
V &= F_Y(y) \\
&= \int_{-\infty}^{x} \sum_{l=1}^{3} \pi_l \times N(t; \lambda_l, \tau_l^2) dt \\
&= \pi_1 \int_{-\infty}^{x} N(t; \lambda_1, \tau_1^2) dt + \pi_2 \int_{-\infty}^{x} N(t; \lambda_2, \tau_2^2) dt + \pi_3 \int_{-\infty}^{x} N(t; \lambda_3, \tau_3^2) dt \\
&\sim U(0,1)
\end{aligned}
$$

4.3.2 Using copulas to measure the FC-SC association for within- vs. between-module edges

Topic 2 (Chapter 3) measured the FC-SC association at each edge in the network, and suggested that the FC-SC association for edges located within the same functional module tended to be slightly higher than for edges between different functional modules. Since within-module edges tend to have both high SC and FC (see Figure 3.5), this result was expected, but the difference was not as evident as we hypothesized, perhaps due to the small sample size or the limitations of the linear link function.

We aim to compare the FC-SC association for within- vs. between-module edges, across all control subject. To perform this analysis, we will concatenate the edge-level connectivity data across subjects, separately for the within- and between-module edge groups, and fit GM models to the within- vs between-module FC and SC marginal distributions. (By collapsing the data across subjects, we make the assumption that the distributions of SC and FC values is the same across all edges.) We will separately measure the FC-SC association for these two classes of edges using Pearson's correlation and 5 different copula functions (as outlined in section 4.1.1). We expect our analysis to show that within-module edges have a higher level of FC-SC association, relative to between-module edges.

4.3.3 Using copulas to measure the edgewise FC-SC association

Copulas provide a flexible way to quantify the dependence between FC and SC, since they can capture linear and non-linear associations, unlike Pearson's correlation or the linear link function described in Topic 2. With a sufficiently large number of subjects, one could use copulas to fit the edgewise SC and FC marginal distributions, and model the FC-SC association at the edge-level, in an analogous way to the modeling approach from Topic 2. Due to the small size of our rs-fMRI and DTI dataset (n=20 control subjects) however, it is not feasible to fit a GM model to the connectivity data for each edge individually. Rather, we make the simplifying assumption that the SC and FC data follows the same distribution for each edge in the network, and concatenate the edge-level connectivity data across all subjects before estimating the distribution of SC and FC with a GM model.

4.4 Simulation Studies

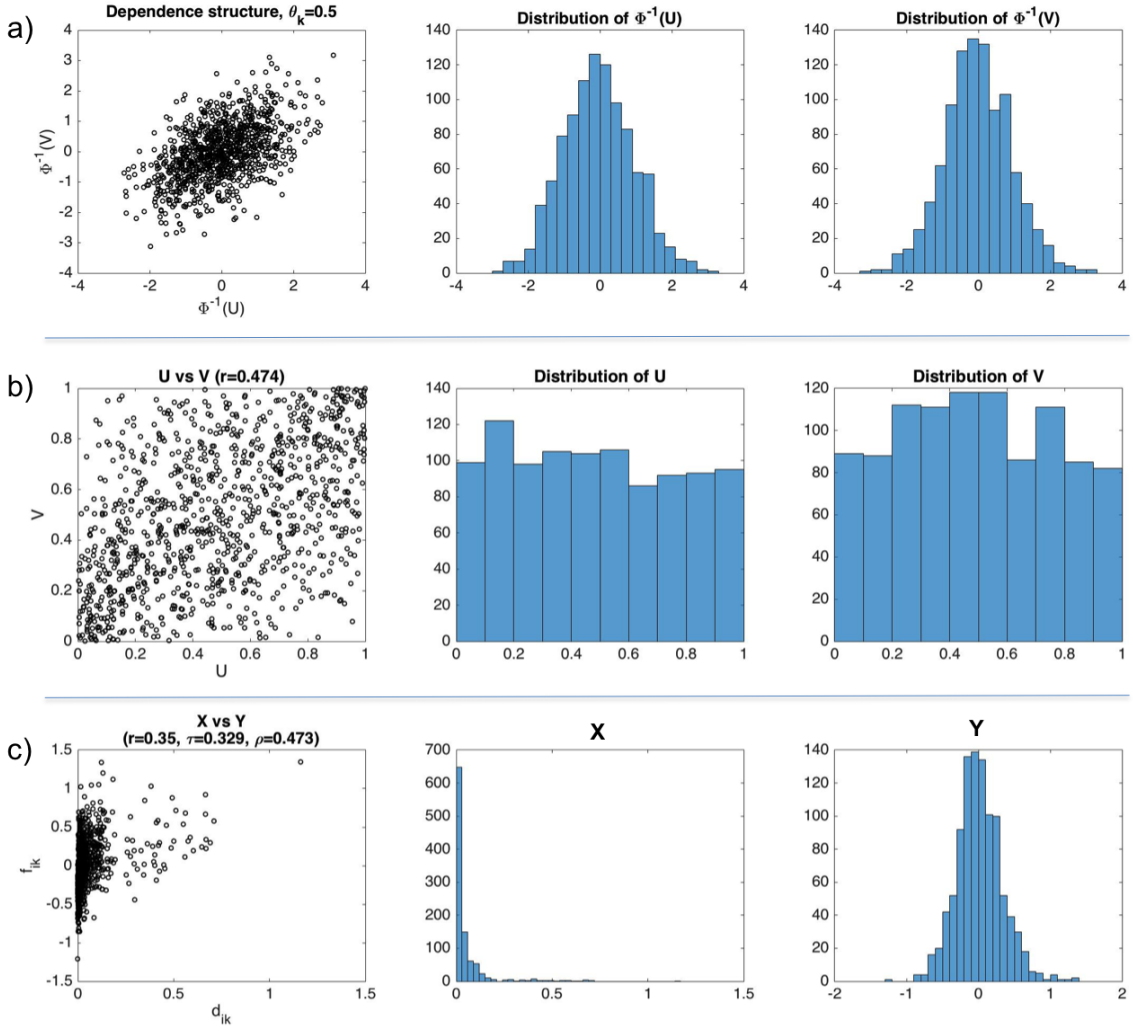
We perform simulation studies to evaluate the estimation method, and to compare how well the copula measures the FC-SC association in the context of our modeling framework, relative to Pearson's correlation. The copula approach is extremely useful for simulating multimodal neuroimaging data, since it allows the user to control the type and level of dependency between modalities, while simultaneously allowing for any choice of marginal distribution for each individual modality (Silva et al., 2014).

4.4.1 Simulation example

Figure 4.5 shows an example of how the FC and SC data can be simulated using the copula framework. The first step in the simulation is to generate data with a known dependence structure specified the copula function and parameter θ. In row a) of Figure 4.5, 1000 data points are generated using a Gaussian copula with $\theta = 0.5$ (corresponding to a Kendall's τ of 0.33). Individually, the random variables $\Phi^{-1}(u)$ and $\Phi^{-1}(v)$ are marginally normally distributed. In row b) of the figure, we calculate the uniformly distributed marginal CDFs using $u = \Phi(\Phi^{-1}(u))$ and $v = \Phi(\Phi^{-1}(v))$. Although the random variables U and V have marginal uniform distributions, they still maintain a strong bivariate association (corr$(U, V) = 0.474$). Row c) shows the SC and FC data (X and Y), distributed according to the user-defined marginal PDFs (4- and 3-component GM models, respectively). In this case, the strong bivariate association specified by the Gaussian copula with $\theta = 0.5$ is maintained, as evidenced the Spearman's $\rho(X, Y) = 0.473$, although the Pearson's linear correlation coefficient is r$(X, Y) = 0.35$.

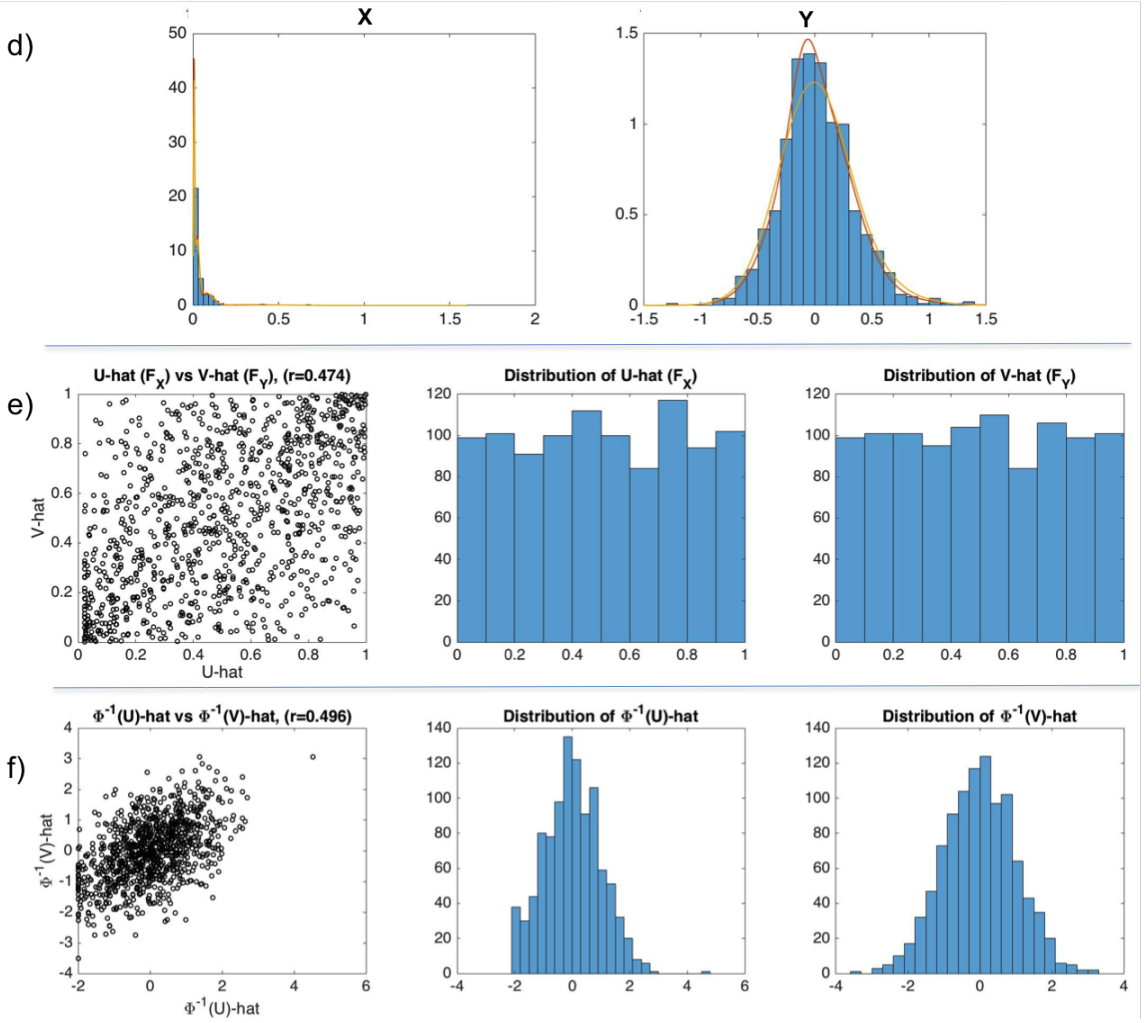
Once the FC and SC data have been simulated with the specified dependence structure and marginal distributions, we can estimate the FC-SC association us-

Figure 4.5: Steps for simulating data under a specified dependence structure



ing the copula, and compare to the ground truth. The first step is to estimate the marginal distributions of X and Y based on the data. Row d) of Figure 4.6 shows the histograms of the simulated data (X and Y), along with the true and fitted GM model PDFs (in red and yellow, respectively). Row e) shows that the estimated marginal CDFs of X and Y, $F_X(x)$ and $F_Y(y)$ are each uniformly distributed, and maintain a strong bivariate association (corr$(\hat{U}, \hat{V}) = 0.474$). In row f) we estimate the FC-SC association using the Gaussian copula: $\hat{\theta} = \text{corr}[\Phi^{-1}(\hat{u}), \Phi^{-1}(\hat{v})] = 0.496$, which is close to the true dependence parameter of $\theta = 0.5$.

Figure 4.6: Estimation steps, based on simulated data



4.4.2 Simulation Results

Following the example outlined in section 4.4.1, we simulate FC and SC data (X, Y) using the Gaussian copula with a range of dependence levels ($\theta = 0, 0.1, 0.2, 0.3, 0.4, 0.5$) and sample sizes ($n = 50, 100, 200$). The parameters for the marginal GM models are specified according to our empirical observations, to generate realistic distributions of SC and FC data. For each condition, we perform 500 simulation runs, and 200 bootstrap resamples of size n per run, to estimate the bootstrap standard error (SE_B) of $\hat{\theta}$. We can estimate the coverage probabilities for the 95% confidence inter-

vals based on 1) the Wald-type interval using $SE_B(\hat{\theta})$ and 2) the bootstrap sample quantiles.

These results are displayed in Table 4.2. We see that under all conditions of dependence level θ and sample size n, we obtain parameter estimates with very low bias. SE_B provides an accurate measure of variability, as compared with the empirical standard deviation of the mean ML estimates. Both CI types achieve close to 95% coverage.

Table 4.2: Simulation results for θ (S=500, B=200)

θ	n	$\hat{\theta}$		SE_B	95% CI	95% CI
		mean	(SD)		cov prob I	cov prob II
0	50	0.004	(0.141)	0.141	94.2	95.8
	100	0.001	(0.098)	0.098	94	95.8
	200	0	(0.072)	0.070	95	94.6
0.1	50	0.082	(0.147)	0.14	92.8	94.6
	100	0.102	(0.101)	0.097	93	94.2
	200	0.095	(0.070)	0.069	93.6	93
0.2	50	0.188	(0.143)	0.135	91.2	93.6
	100	0.194	(0.096)	0.095	93.8	94.4
	200	0.193	(0.070)	0.067	93.6	93.2
0.3	50	0.282	(0.138)	0.131	91.6	93.6
	100	0.293	(0.091)	0.091	94.2	95
	200	0.292	(0.065)	0.064	92.8	94.4
0.4	50	0.368	(0.128)	0.126	92.8	94
	100	0.392	(0.085)	0.086	95.4	95.8
	200	0.401	(0.063)	0.059	92.8	93.8
0.5	50	0.468	(0.116)	0.118	93	95.2
	100	0.488	(0.077)	0.080	94.6	93.8
	200	0.495	(0.054)	0.054	94.2	93.8

Tables 4.3 and 4.4 show the sensitivity and specificity of the Gaussian copula dependence parameter (θ) relative to other measures of the X-Y association, namely Pearson's correlation (r), Kendall's tau (τ), and Spearman's rho (ρ). In general, we see that sensitivity is highest for the Gaussian copula dependence parameter, especially when compared to Pearson's correlation. As expected, sensitivity for

each of the association measures increases as the sample size (n)and true association value (θ) increase. When the true FC-SC association is very high (e.g. $\theta = 0.5$), there is no distinction in sensitivity between the different measures of association. Interestingly, the Gaussian copula dependence parameter shows no improvement in specificity over the other measures when $\theta = 0$ (see Figure 4.4).

Table 4.3: Simulation results, Sensitivity(%)

θ	n	$\hat{\theta}$	r	τ	ρ
0.1	50	**13**	9.4	8.4	9.2
	100	**20.2**	12	16.6	16.6
	200	**29.6**	16	24	23.6
0.2	50	**30.6**	17	25.2	27.4
	100	**52**	30.4	48.8	49.2
	200	**79.8**	49.4	76.2	76.8
0.3	50	**57**	32.8	49	51
	100	**88.2**	56.6	84.4	85
	200	**98.6**	83.8	97.6	97.6
0.4	50	**81.6**	52.4	75	77.4
	100	**98.2**	83	98	97.8
	200	100	99	100	100
0.5	50	93.6	77.2	94.2	94.2
	100	100	97.6	100	100
	200	100	99.8	100	100

Table 4.4: Simulation results, Specificity (%)

θ	n	$\hat{\theta}$	r	τ	ρ
0	50	94.2	95.6	96.4	96.2
	100	94	96.8	95	95.2
	200	95	96	97	97.2

The sensitivity results from Table 4.3 are displayed graphically in Figure 4.7. When visualized in this way, the improvement in sensitivity of the Gaussian copula dependence parameter relative to Pearson's linear correlation coefficient is very evident. This suggests that linear correlation may not be the best measure of edgewise FC-SC association. The rank correlation coefficients, which have a

one-to-one mapping with the Gaussian copula dependence parameter θ, also display an improvement in sensitivity relative to Pearson's correlation. However, the copula approach has the advantage in that it can also be used to specify the joint distribution of FC and SC, in addition to providing a scalar measure of association.

4.5 Data Analysis

We apply the proposed copula method for the 20 healthy control subjects. The 232 $\times$ 232 mean SC and FC matrices are displayed in Figure 3.5 a)-b). For this analysis, we only consider full correlation as a measure of FC. As expected, we observe that SC and FC are both high for within-module edges, and low for between-module edges.

4.5.1 Using copulas to measure the FC-SC association for within- vs. between-module edges

To compare the FC-SC association for within- vs. between-module edges, we first fit GM models to each subject's marginal SC and FC data. Figure 4.8 shows the SC data from both edge groups for a randomly chosen control subject (con26), overlaid with the pdf of the fitted 4-component GM model (in orange). The GM models appear to provide a reasonable fit for both edge groups. The histograms show that SC tends to be higher for within-module edges. The mean and standard deviation of the GM model parameters estimated across all subjects' SC' data are shown in Tables 4.5 and 4.6.

Figure 4.9 demonstrates that we can reasonably fit a GM model to the subject-level within- and between-module FC' data, as shown using one subject's data (con26). The fitted GM model pdf is overlaid (orange line) on the histogram of observed data. Since FC tends to be positive for within-module edges, a 2-component

Figure 4.7: Sensitivity of different measures of dependence

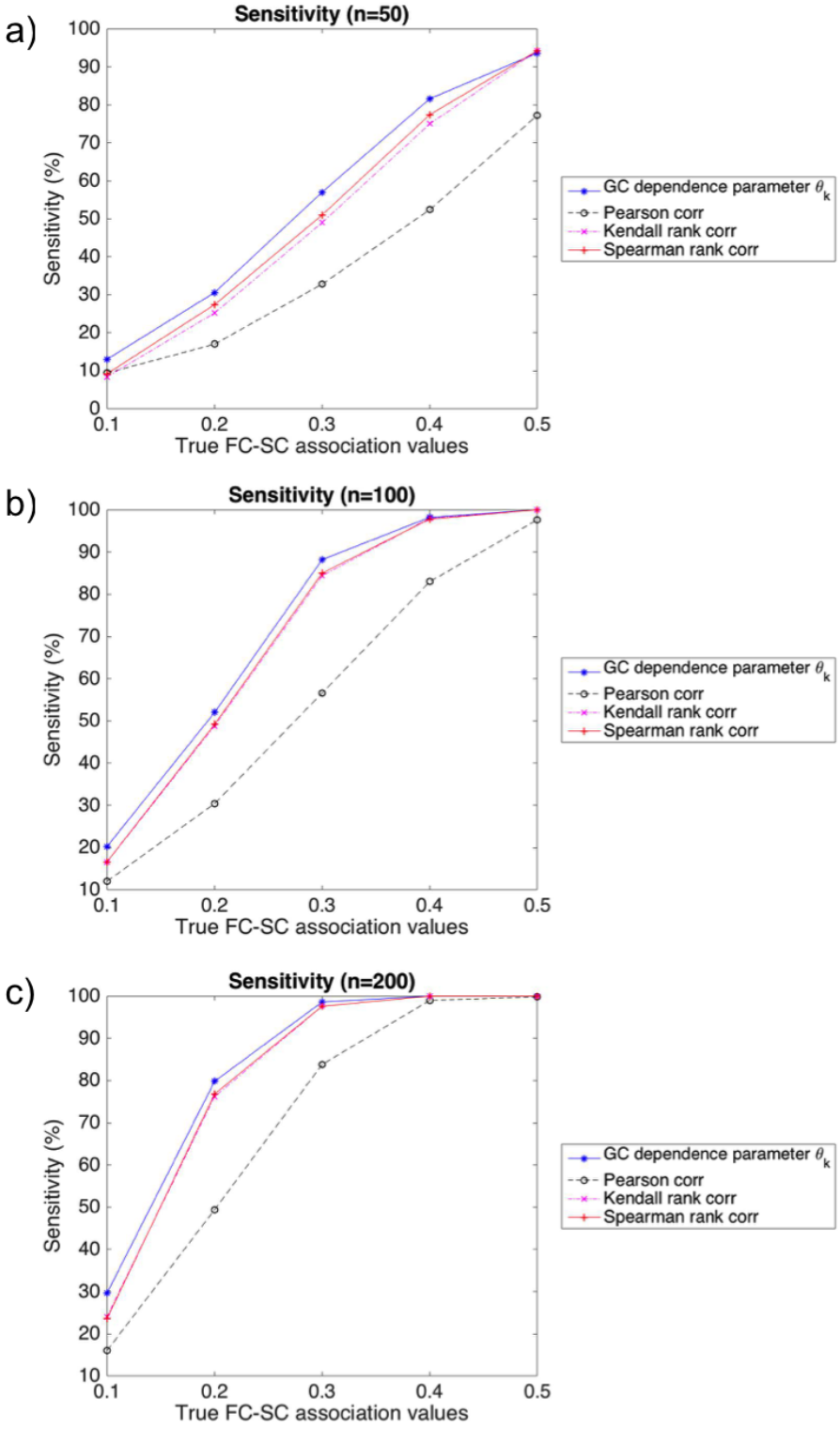


Figure 4.8: GM models fitted to SC data, at within- vs. between-module edges

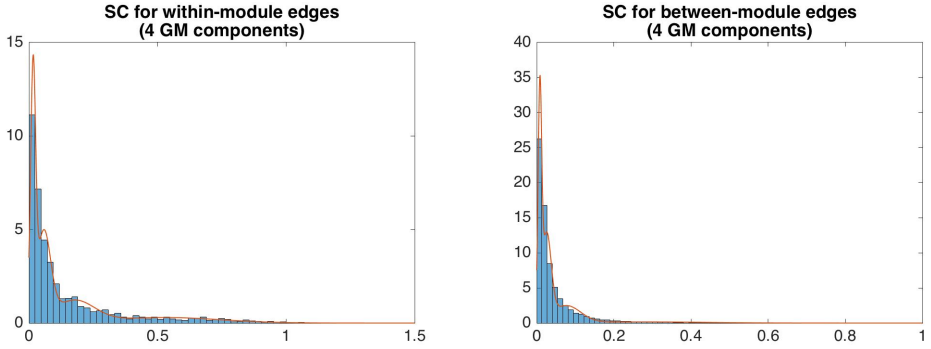


Table 4.5: GM model components extracted from within-module SC′ data

component	μ mean (SD)	σ mean (SD)	p mean (SD)
near-zero SC	0.010 (0.004)	0.006 (0.002)	0.24 (0.05)
very low SC	0.040 (0.01)	0.019 (0.004)	0.31 (0.03)
mod. SC	0.131 (0.021)	0.063 (0.008)	0.26 (0.03)
high SC	0.473 (0.026)	0.228 (0.005)	0.19 (0.02)

Table 4.6: GM model components extracted from between-module SC′ data

component	μ mean (SD)	σ mean (SD)	p mean (SD)
near-zero SC	0.006 (0.002)	0.003 (0.001)	0.37 (0.03)
very low SC	0.021 (0.005)	0.010 (0.002)	0.35 (0.02)
mod. SC	0.068 (0.011)	0.035 (0.004)	0.22 (0.01)
high SC	0.259 (0.017)	0.169 (0.004)	0.07 (0.01)

GM model, representing the positive and near-zero FC elements, provides a better fit to this data. The between-module edges can be fit with a 3-component GM model, as proposed. The mean and standard deviation of the GM model parameters estimated across all subjects' FC′ data are shown in Tables 4.7 and 4.8.

Figures 4.10 and 4.11 show the FC-SC association for within- and between-module edges, respectively, at the level of the original variables X, Y (*on left*) and their marginal uniform CDFs U, V (*on right*). (The data shown is for one control subject, con34.) In each case, the plots of U vs V don't necessarily resemble the

Figure 4.9: GM models fitted to FC data, at within- vs. between-module edges

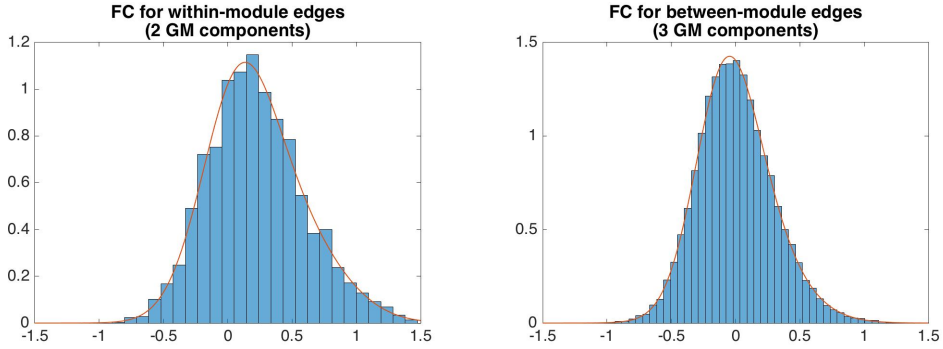


Table 4.7: GM model components extracted from FC′ data

component	within-module		
	μ	σ	p
	mean (SD)	mean (SD)	mean (SD)
negative FC	-	-	-
near-zero FC	0.074 (0.06)	0.265 (0.04)	0.54 (0.13)
positive FC	0.372 (0.09)	0.364 (0.03)	0.46 (0.13)

Table 4.8: GM model components extracted from FC′ data

component	between-module		
	μ	σ	p
	mean (SD)	mean (SD)	mean (SD)
negative FC	-0.145 (0.07)	0.222 (0.04)	0.36 (0.10)
near-zero	0.046 (0.05)	0.234 (0.03)	0.41 (0.10)
positive FC	0.141 (0.06)	0.346 (0.05)	0.24 (0.12)

dependence structures shown by any of the five copula functions (see Figures 4.2 and C.1-C.5), but it is difficult to discern because the association is low.

Table 4.9 shows the FC-SC association as measured by correlation coefficients between the random variables X and Y, as well as the p-value for the Wilcoxon rank sum test comparing the two edge groups. The FC-SC association is significantly stronger for within-module edges, compared to between-module edges, as measured by each of the three correlation coefficients.

Likewise, Table 4.10 shows the FC-SC association as measured by the copula

Figure 4.10: FC-SC association for within-module edges

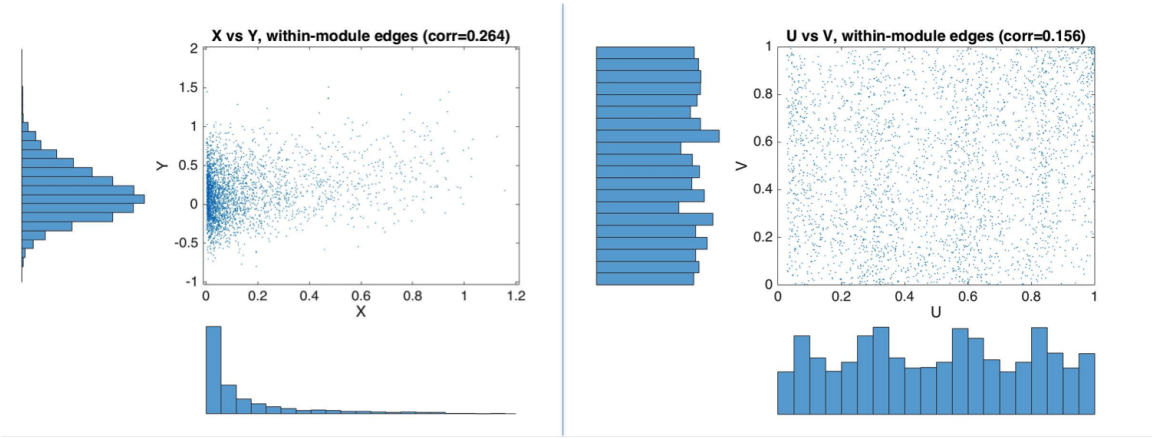
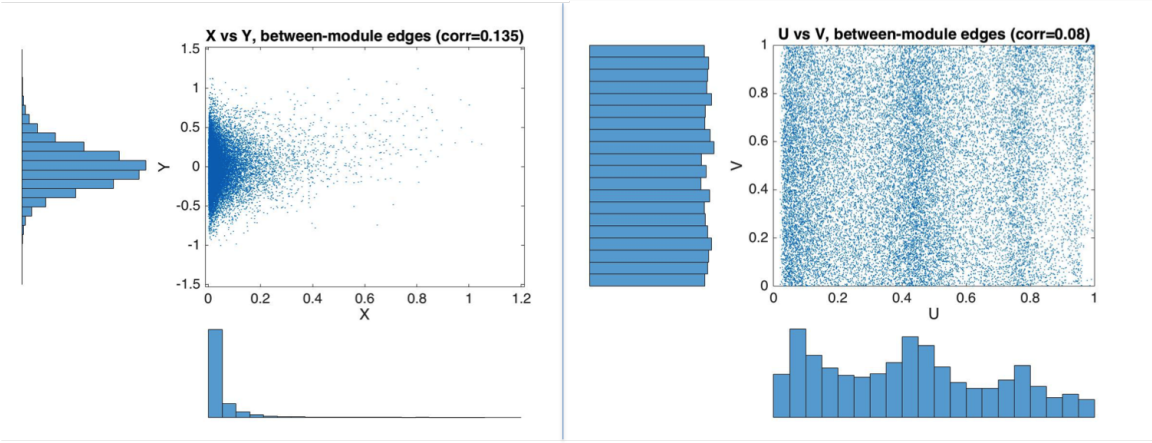


Figure 4.11: FC-SC association for between-module edges



dependence parameter and its corresponding Kendall's τ, for five different copula functions. In each case, the measured FC-SC association is significantly higher for within-module edges, as we hypothesized, except when measured by the Clayton copula ($p = 0.232$).

4.5.2 Using copulas to measure the edgewise FC-SC association

In addition to using copulas to compare the FC-SC association for within- vs between-module edges, we also aimed to use them to estimate the edgewise FC-SC associ-

Table 4.9: FC-SC association for within- vs. between-module edges, as measured correlation coefficients

	within-module edges mean (SD)	between-module edges mean (SD)	p
Pearson's corr	0.272 (0.04)	0.132 (0.02)	6.8×10^{-8}
Kendall's tau	0.096 (0.03)	0.033 (0.02)	1.05×10^{-6}
Spearman's rho	0.143 (0.05)	0.050 (0.03)	1.38×10^{-6}

Table 4.10: FC-SC association for within- vs. between-module edges, as measured by five copula functions

copula type	within-module edges		between-module edges		p
	θ mean (SD)	τ mean (SD)	θ mean (SD)	τ mean (SD)	
Gaussian	0.168 (0.05)	0.107 (0.03)	0.073 (0.03)	0.046 (0.02)	1.20×10^{-6}
t	0.167 (0.06)	0.107 (0.04)	0.065 (0.03)	0.042 (0.20)	1.05×10^{-6}
Clayton	0.062 (0.06)	0.115 (0.03)	0.033 (0.03)	0.016 (0.01)	0.232
Gumbel	1.131 (0.03)	0.115 (0.03)	1.060 (0.03)	0.057 (0.01)	1.23×10^{-7}
Frank	0.900 (0.30)	0.099(0.03)	0.333 (0.16)	0.037 (0.02)	6.01×10^{-7}

ation, at each edge in the whole-brain network. Figures 4.12 and 4.13 show the unthresholded heatmaps of the estimated Clayton and Frank copula dependence parameter θ (*on left*) and associated Kendall's τ (*on right*) across the whole brain network. The Clayton and Frank copulas yield results that seem to be biologically meaningful, since they tend to show high levels of FC-SC association for within-module edges. The results for the other three fitted copula functions are displayed in Figures C.6-C.8. It should be noted, however, that the Clayton and Gumbel copula functions can only measure positive associations, which makes them more limited compared to the Gaussian, t, and Frank copula functions, which can measure both positive and negative associations. Also, we note although the association parameter θ has different ranges for different copulas, it can be expressed in terms of Kendall's τ, which is comparable between the different copula functions. The edges with significant FC-SC associations, as measured by the Clayton and Frank copulas, are shown by the thresholded heatmaps in Figure 4.14 (uncorrected

p<0.001).

Figure 4.12: Edgewise FC-SC association, measured by the Clayton copula

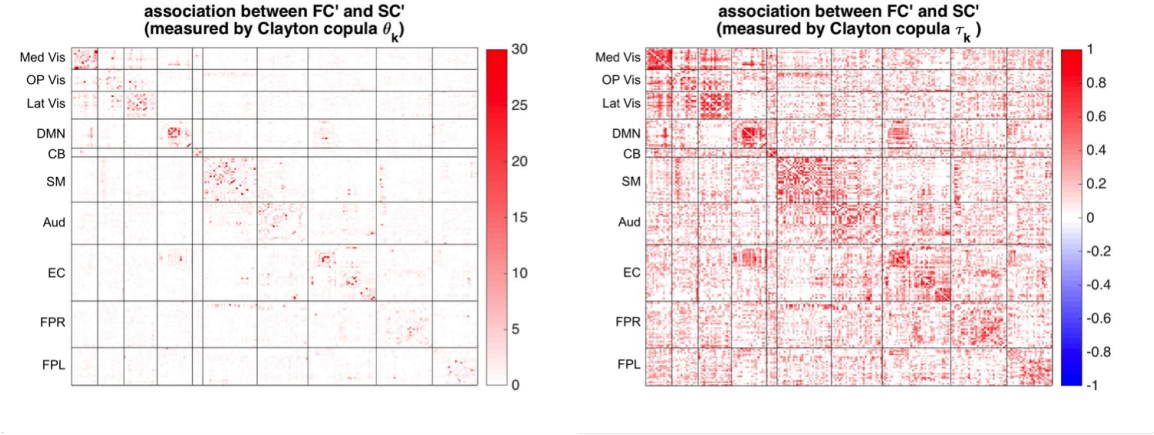
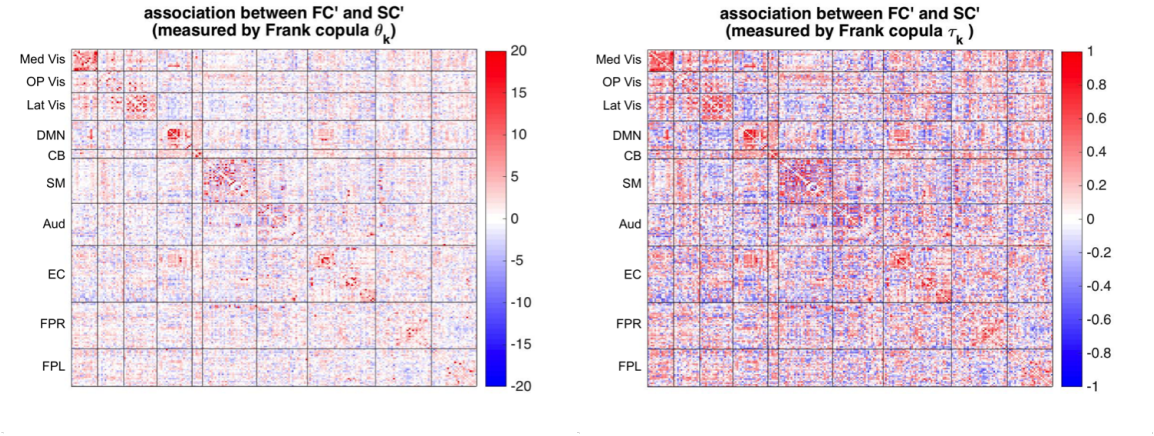


Figure 4.13: Edgewise FC-SC association, measured by the Frank copula



In addition, if we measure FC using partial correlations rather than full correlations, we can show that the copula-based association measure τ allows for direct comparison of the results. This is because the copula association measure does not depend on the marginal behavior of the univariate connectivity measures. For example, Figure 4.15 shows the heatmap of the Frank copula-based association measure (τ_k), when FC is measured by full vs. partial correlations, to show that the results are on the same scale.

Figure 4.14: Thresholded edgewise FC-SC association (p<0.001)

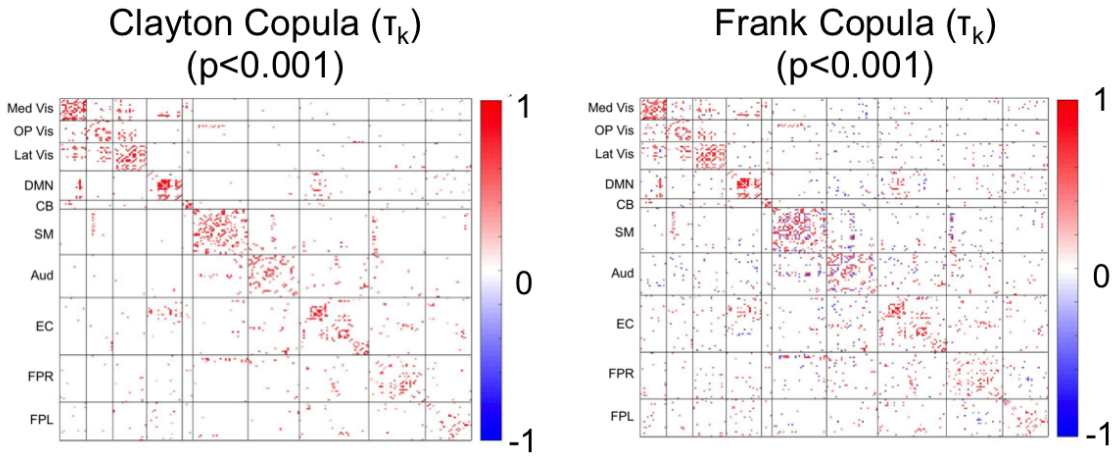
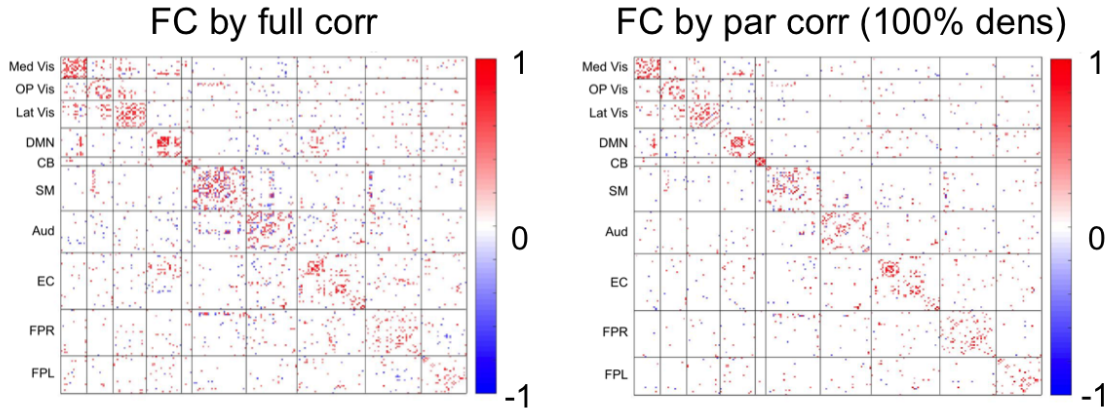


Figure 4.15: Frank copula-based edgewise FC-SC association (p<0.001), for FC measured in two ways



4.6 Discussion

Copula functions allow us to simultaneously describe the dependence structure between random variables, and separately, to model their marginal univariate distributions. It is a useful tool for measuring the association between FC and SC that doesn't assume that their relationship is linear. This method overcomes the limitations of Pearson's linear correlation and the linear link function in Topic 2,

by offering a more flexible way to model the FC-SC association. Compared to the method in Topic 2, the copula-based method provides a more robust measure of association that does not depend on the marginal behavior of FC and SC values, and can be directly compared between different measures.

Copulas are especially useful for simulating multimodal neuroimaging data under various dependence structures and levels, and with the user's choice of marginal distributions. Our simulation studies found that copulas might offer a more efficient method for detecting FC-SC associations in the brain network. Our simulations only consider the dependence structure prescribed by the Gaussian copula function. Future simulation studies will test the performance of other copula functions.

Our data analysis with an rs-fMRI and DTI dataset of 20 healthy controls found that the FC-SC association, as measured by five different types of copulas, is higher for within-module edges compared to between-module edges, as we hypothesized. Our edgewise copula analysis with the Frank and Clayton copulas also yielded biologically meaningful results, indicating that these might be good candidate copulas for measuring multimodal relationships in the brain, although the Clayton copula can only model positive associations. This analysis confirms that copulas can be successfully applied in providing additional important information about the function-structure associations in the brain network, that cannot be provided by Pearson's correlation coefficient.

Future analyses will consider partial correlations as a measure of FC, and will also perform group comparisons between the MDD subjects and healthy controls. Also, we caution that the results of our analysis are based on a small data set of 20 healthy control subjects. A more rigorous validation should be performed, based on a larger sample of subjects.

Appendix A

Appendix for Chapter 2 (Topic 1)

The numerator and denominator of $\hat{\theta}_\ell$ can be written as linear combinations of $\boldsymbol{N^*}$

$$\hat{\theta}_\ell = \frac{(\boldsymbol{C_\ell} - \boldsymbol{A})\boldsymbol{N^*}}{b - \boldsymbol{A}\boldsymbol{N^*}}$$

$$\text{where} \quad \boldsymbol{A} = \frac{(V_\ell - 1)}{2(V-1)} \sum_{j \in \Omega_\ell} \boldsymbol{C_j} \quad \text{and} \quad b = \frac{V_\ell(V_\ell - 1)}{2} N$$

Notation: $\boldsymbol{C_\ell}$ and $\boldsymbol{C_j}$ are vectors of binary indicators, indicating which voxel pairs are members of Ω_ℓ or involve voxel j, respectively. V is the total number of voxels in the whole brain, and V_ℓ is the number of voxels in component ℓ. N is the maximum possible number of connections between a pair of voxels (i.e. the number of streams initiated from each voxel in a probabilistic tractography procedure).

Proof:

$$\hat{\theta}_\ell = \frac{\underbrace{\sum_{j,k \in \Omega_\ell} [N_{jk} - (\bar{N}_j + \bar{N}_k)/2]}_{①}}{\underbrace{\sum_{j,k \in \Omega_\ell} [N - (\bar{N}_j + \bar{N}_k)/2]}_{②}}$$

$\hat{\theta}_\ell$ **numerator:**

$$\begin{aligned}
\textcircled{1} &= \sum_{j,k\in\Omega_\ell} [N_{jk} - (\bar{N}_j + \bar{N}_k)/2] \\
&= \underbrace{\sum_{j,k\in\Omega_\ell} N_{jk}}_{\textcircled{3}} - \underbrace{\sum_{j,k\in\Omega_\ell} [(\bar{N}_j + \bar{N}_k)/2]}_{\textcircled{4}}
\end{aligned}$$

$$\textcircled{3} = \sum_{j,k\in\Omega_\ell} N_{jk} = \begin{bmatrix} 1 & 1 & 0 & 1 & \dots & 0 \end{bmatrix} \begin{bmatrix} N_{12} \\ N_{13} \\ \vdots \\ N_{V-1,V} \end{bmatrix} = \boldsymbol{C_\ell N^*}$$

Where $\boldsymbol{C_\ell}$ is a $1 \times \binom{V}{2}$ vector of binary indicators, indicating which voxel pairs are members of Ω_ℓ.

$$\begin{aligned}
\textcircled{4} &= \sum_{j,k\in\Omega_\ell} [(\bar{N}_j + \bar{N}_k)/2] \\
&= \frac{1}{2} \sum_{j,k\in\Omega_\ell} \left[\frac{[N_{j1} + N_{j2} + \ldots + N_{jV}]}{V-1} + \frac{[N_{k1} + N_{k2} + \ldots + N_{kV}]}{V-1} \right] \\
&= \frac{1}{2} \sum_{j,k\in\Omega_\ell} \left[\frac{\boldsymbol{C_j N^*}}{V-1} + \frac{\boldsymbol{C_k N^*}}{V-1} \right] \\
&= \frac{1}{2(V-1)} \sum_{j,k\in\Omega_\ell} [\boldsymbol{C_j} + \boldsymbol{C_k}] \boldsymbol{N^*} \\
&= \frac{V_\ell - 1}{2(V-1)} \sum_{j\in\Omega_\ell} \boldsymbol{C_j N^*}
\end{aligned}$$

where $\boldsymbol{C_j}$ is a $1 \times \binom{V}{2}$ vector of binary indicators, indicating which voxel pairs include voxel j

$$\therefore \textcircled{1} = \left(\boldsymbol{C_\ell} - \underbrace{\frac{V_\ell - 1}{2(V-1)} \sum_{j \in \Omega_\ell} \boldsymbol{C_j}}_{\boldsymbol{A}} \right) \boldsymbol{N^*} = (\boldsymbol{C_\ell} - \boldsymbol{A}) \boldsymbol{N^*}$$

$\hat{\theta}_\ell$ denominator:

$$\begin{aligned} \textcircled{2} &= \sum_{j,k \in \Omega_\ell} [N - (\bar{N}_j + \bar{N}_k)/2] \\ &= \frac{V_\ell (V_\ell - 1)}{2} N - \underbrace{\sum_{j,k \in \Omega_\ell} [(\bar{N}_j + \bar{N}_k)/2]}_{\textcircled{4}} \end{aligned}$$

$$\therefore \textcircled{2} = \underbrace{\frac{V_\ell (V_\ell - 1)}{2} N}_{b} - \underbrace{\left(\frac{V_\ell - 1}{2(V-1)} \sum_{j \in \Omega_\ell} \boldsymbol{C_j} \right)}_{\boldsymbol{A}} \boldsymbol{N^*} = b - \boldsymbol{A} \boldsymbol{N^*}$$

Appendix B

Appendix for Chapter 3 (Topic 2)

E-step:

$$
\begin{aligned}
Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)}) &= E_{s_{ik}|\boldsymbol{Y},\boldsymbol{\theta}^{(t)}}[\ell(\boldsymbol{\theta}|\boldsymbol{Y}, Z)] \\
&= -\left(\frac{J\lambda_k^2}{2\sigma_k^2} + \frac{L}{2\gamma_k^2} + \frac{1}{2\tau_k^2}\right) \sum_{i=1}^{n} \overbrace{E(s_{ik}^2|\boldsymbol{Y}, \boldsymbol{\theta}^{(t)})}^{E_{i2}} \\
&\quad + \sum_{i=1}^{n} \underbrace{E(s_{ik}|\boldsymbol{Y}, \boldsymbol{\theta}^{(t)})}_{E_{i1}} \left[\frac{1}{\sigma_k^2} \sum_{j=1}^{J} \lambda_k(\tilde{r}_{ik} - \beta_k) + \frac{1}{\gamma_k^2} \sum_{l=1}^{L} p'_{ikl} + \frac{s_k}{\tau_k^2}\right] \\
&\quad + \text{const}
\end{aligned}
$$

where

$$
\begin{aligned}
\text{const} = &-nJ\log(\sqrt{2\pi}\sigma_k) - nL\log(\sqrt{2\pi}\gamma_k) - n\log(\sqrt{2\pi}\tau_k) \\
&- \frac{1}{2\sigma_k^2} \sum_{i=1}^{n}\sum_{j=1}^{J} (\tilde{r}_{ikj}^2 + \beta_k^2 - 2\tilde{r}_{ikj}\beta_k) - \frac{1}{2\gamma_k^2} \sum_{i=1}^{n}\sum_{l=1}^{L} p'^2_{ikl} - \frac{ns_k^2}{2\tau_k^2}
\end{aligned}
$$

and

$$E_{i2} = E(s_{ik}^2|\boldsymbol{Y}, \boldsymbol{\theta}^{(t)}) = [E(s_{ik}|\boldsymbol{Y}, \boldsymbol{\theta}^{(t)})]^2 + Var(s_{ik}|\boldsymbol{Y}, \boldsymbol{\theta}^{(t)})$$

To find E_{i1} and E_{i2}, we need $f(s_{ik}|\boldsymbol{Y}, \boldsymbol{\theta}^{(t)}))$

$$\begin{aligned} f(s_{ik}|\boldsymbol{Y}, \boldsymbol{\theta}^{(t)}) &= \frac{f(\boldsymbol{Y}, s_{ik}|\boldsymbol{\theta}^{(t)})}{f(\boldsymbol{Y}|\boldsymbol{\theta}^{(t)})} \\ &\propto f(\boldsymbol{Y}, s_{ik}|\boldsymbol{\theta}^{(t)}) = L(\boldsymbol{\theta}^{(t)}|\boldsymbol{Y}, s_{ik}) \end{aligned}$$

$$\begin{aligned} f(s_{ik}|\boldsymbol{Y}, \boldsymbol{\theta}^{(t)}) \propto \prod_{i=1}^{n} \Bigg[\prod_{j} \frac{1}{\sqrt{2\pi}\sigma_k} e^{\left(-\frac{[\tilde{r}_{ikj} - (\lambda_k s_{ik} + \beta_k)]^2}{2\sigma_k^2}\right)} \\ \times \prod_{l} \frac{1}{\sqrt{2\pi}\gamma_k} e^{\left(-\frac{(p'_{ikl} - s_{ik})^2}{2\gamma_k^2}\right)} \times \frac{1}{\sqrt{2\pi}\tau_k} e^{\left(-\frac{(s_{ik} - s_k)^2}{2\tau_k^2}\right)} \Bigg] \end{aligned}$$

For a given i and k we can write

$$= \frac{e^{-\sum_{j=1}^{J} \frac{(\tilde{r}_{ikj} - \lambda_k s_{ik} - \beta_k)^2}{2\sigma_k^2}}}{(\sqrt{2\pi}\sigma_k)^J} + \frac{e^{-\sum_{l=1}^{L} \frac{(p'_{ikl} - s_{ik})^2}{2\gamma_k^2}}}{(\sqrt{2\pi}\gamma_k)^L} + \frac{e^{-\frac{(s_{ik} - s_k)^2}{2\tau_k^2}}}{\sqrt{2\pi}\tau_k}$$

The negative exponent term in the numerator boils down to...

$$\begin{aligned} &\sum_{j=1}^{J} \frac{(\tilde{r}_{ikj} - \lambda_k s_{ik} - \beta_k)^2}{2\sigma_k^2} + \sum_{l=1}^{L} \frac{(p'_{ikl} - s_{ik})^2}{2\gamma_k^2} + \frac{(s_{ik} - s_k)^2}{2\tau_k^2} \\ &= \sum_{j=1}^{J} \frac{(\tilde{r}_{ikj} - \beta_k)^2 - 2\lambda_k s_{ik}(\tilde{r}_{ikj} - \beta_k) + \lambda_k^2 s_{ik}^2)}{2\sigma_k^2} + \sum_{l=1}^{L} \frac{(p'^2_{ikl} - 2p'_{ikl} s_{ik} + s_{ik}^2)}{2\gamma_k^2} \end{aligned}$$

$$+\frac{(s_{ik}^2-2s_{ik}s_k+s_k^2)}{2\tau_k^2}$$

$$=\sum_{j=1}^{J}\frac{(\lambda_k^2 s_{ik}^2-2\lambda_k(\tilde{r}_{ikj}-\beta_k)s_{ik})}{2\sigma_k^2}+\sum_{l=1}^{L}\frac{(s_{ik}^2-2p'_{ikl}s_{ik})}{2\gamma_k^2}+\frac{(s_{ik}^2-2s_{ik}s_k)}{2\tau_k^2}+C$$

$$=\frac{J\lambda_k^2 s_{ik}^2}{2\sigma_k^2}-\frac{2\lambda_k(\tilde{r}_{ik\bullet}-J\beta_k)s_{ik}}{2\sigma_k^2}+\frac{Ls_{ik}^2}{2\gamma_k^2}-\frac{2p'_{ik\bullet}s_{ik}}{2\gamma_k^2}+\frac{s_{ik}^2}{2\tau_k^2}-\frac{2s_k s_{ik}}{2\tau_k^2}+C$$

$$=s_{ik}^2\left[\frac{J\lambda_k^2}{2\sigma_k^2}+\frac{L}{2\gamma_k^2}+\frac{1}{2\tau_k^2}\right]-2s_{ik}\left[\frac{\lambda_k(\tilde{r}_{ik\bullet}-J\beta_k)}{2\sigma_k^2}+\frac{p'_{ik\bullet}}{2\gamma_k^2}+\frac{s_k}{2\tau_k^2}\right]+C$$

$$=\left[\frac{J\lambda_k^2}{2\sigma_k^2}+\frac{L}{2\gamma_k^2}+\frac{1}{2\tau_k^2}\right]\left(s_{ik}^2-2s_{ik}\frac{\left[\frac{\lambda_k(\tilde{r}_{ik\bullet}-J\beta_k)}{2\sigma_k^2}+\frac{p'_{ik\bullet}}{2\gamma_k^2}+\frac{s_k}{2\tau_k^2}\right]}{\left[\frac{J\lambda_k^2}{2\sigma_k^2}+\frac{L}{2\gamma_k^2}+\frac{1}{2\tau_k^2}\right]}\right)+C$$

$$=\frac{1}{2\sigma_{ik}^{*2}}(s_{ik}-\mu_{ik}^*)^2+C$$

where

$$\sigma_{ik}^{*2}=\frac{1}{\left[\frac{J\lambda_k^2}{\sigma_k^2}+\frac{L}{\gamma_k^2}+\frac{1}{\tau_k^2}\right]}\qquad\mu_{ik}^*=\frac{\left[\frac{\lambda_k(\tilde{r}_{ik\bullet}-J\beta_k)}{\sigma_k^2}+\frac{p'_{ik\bullet}}{\gamma_k^2}+\frac{s_k}{\tau_k^2}\right]}{\left[\frac{J\lambda_k^2}{\sigma_k^2}+\frac{L}{\gamma_k^2}+\frac{1}{\tau_k^2}\right]}$$

Notation: $\tilde{r}_{ik\bullet}=\sum_{j=1}^{J}\tilde{r}_{ikj}$ and $p'_{ik\bullet}=\sum_{l=1}^{L}p'_{ikl}$

Thus

$[s_{ik}|\boldsymbol{\theta}^{(t)},\boldsymbol{Y}]\sim N(\mu_{ik}^*,\sigma_{ik}^{*2})$

and

$\boxed{E_{i1}=E(s_{ik}|\boldsymbol{Y},\boldsymbol{\theta}^{(t)})=\mu_{ik}^*}$

$\boxed{E_{i2}=E(s_{ik}^2|\boldsymbol{Y},\boldsymbol{\theta}^{(t)})=[E(s_{ik}|\boldsymbol{Y},\boldsymbol{\theta}^{(t)})]^2+Var(s_{ik}|\boldsymbol{Y},\boldsymbol{\theta}^{(t)})=\mu_{ik}^{*2}+\sigma_{ik}^{*2}}$

Table B.1: Simulation results for β_k, n=20

setting	β_k	$\hat{\beta}_k$ mean (SD)	$\hat{SE}_B$	CI cov prob I	CI cov prob II
high SC, high FC	0.7	0.703 (0.053)	0.05	92	92
	1	1 (0.053)	0.05	91	90
	0.7	0.7 (0.052)	0.051	93	91
	1	1.002 (0.052)	0.051	93	92
med SC, med FC	0.7	0.704 (0.086)	0.083	91	91
	1	1.001 (0.086)	0.084	91	90
high SC, low FC	0.2	0.199 (0.053)	0.049	90	89
	0.2	0.2 (0.052)	0.05	92	92
low SC, high FC	0.5	0.502 (0.188)	0.195	93	93
	0.7	0.7 (0.19)	0.194	94	92
low SC, low FC	0.2	0.196 (0.195)	0.195	93	92
	0.2	0.199 (0.194)	0.194	92	91
	0.2	0.188 (0.191)	0.191	93	91
	0.4	0.4 (0.195)	0.195	93	91
	0.4	0.403 (0.193)	0.193	92	91

Table B.2: Simulation results for β_k, n=50

setting	β_k	$\hat{\beta}_k$ mean (SD)	$\hat{SE}_B$	CI cov prob I	CI cov prob II
high SC, high FC	0.7	0.699 (0.031)	0.031	94	94
	1	1 (0.032)	0.031	92	92
	0.7	0.699 (0.032)	0.031	93	93
	1	1.003 (0.031)	0.031	94	94
med SC, med FC	0.7	0.701 (0.051)	0.049	93	92
	1	1.001 (0.050)	0.050	93	93
high SC, low FC	0.2	0.201 (0.031)	0.031	94	93
	0.2	0.2 (0.031)	0.030	94	94
low SC, high FC	0.5	0.502 (0.119)	0.114	94	93
	0.7	0.698 (0.116)	0.115	94	94
low SC, low FC	0.2	0.195 (0.115)	0.115	94	94
	0.2	0.203 (0.117)	0.115	94	93
	0.2	0.203 (0.119)	0.115	94	93
	0.4	0.403 (0.115)	0.117	94	93
	0.4	0.408 (0.114)	0.115	94	94

Table B.3: Simulation results for s_k, n=20

setting	s_k	$\hat{s}_k$ mean (SD)	$\hat{SE}_B$	CI cov prob I	CI cov prob II
high SC, high FC	0.41	0.403 (0.103)	0.098	92	91
	0.41	0.405 (0.103)	0.096	92	92
	0.41	0.402 (0.104)	0.097	93	93
	0.41	0.404 (0.105)	0.098	92	92
med SC, med FC	-0.85	-0.85 (0.107)	0.096	90	90
	-0.85	-0.843 (0.102)	0.097	93	93
high SC, low FC	0.41	0.408 (0.105)	0.096	93	93
	0.41	0.403 (0.1)	0.097	93	93
low SC, high FC	-2.20	-2.204 (0.1)	0.097	93	92
	-2.20	-2.197 (0.103)	0.098	92	92
low SC, low FC	-2.20	-2.196 (0.102)	0.097	92	92
	-2.20	-2.189 (0.099)	0.096	93	93
	-2.20	-2.203 (0.098)	0.098	93	93
	-2.20	-2.2 (0.097)	0.096	94	94
	-2.20	-2.195 (0.101)	0.097	93	92

Table B.4: Simulation results for s_k, n=50

setting	s_k	$\hat{s}_k$ mean (SD)	$\hat{SE}_B$	CI cov prob I	CI cov prob II
high SC, high FC	0.41	0.404 (0.066)	0.063	94	93
	0.41	0.408 (0.065)	0.063	94	94
	0.41	0.407 (0.064)	0.063	94	93
	0.41	0.404 (0.066)	0.063	93	93
med SC, med FC	-0.85	-0.847 (0.065)	0.063	93	93
	-0.85	-0.846 (0.062)	0.063	94	94
high SC, low FC	0.41	0.407 (0.067)	0.063	92	91
	0.41	0.404 (0.065)	0.063	94	93
low SC, high FC	-2.20	-2.196 (0.066)	0.063	93	92
	-2.20	-2.196 (0.062)	0.063	95	95
low SC, low FC	-2.20	-2.196 (0.065)	0.063	94	94
	-2.20	-2.202 (0.064)	0.063	95	94
	-2.20	-2.202 (0.065)	0.063	94	93
	-2.20	-2.197 (0.064)	0.063	95	93
	-2.20	-2.197 (0.064)	0.063	94	93

Figure B.1: Significant standardized $\hat{\lambda}_k$ results, using different measures of FC

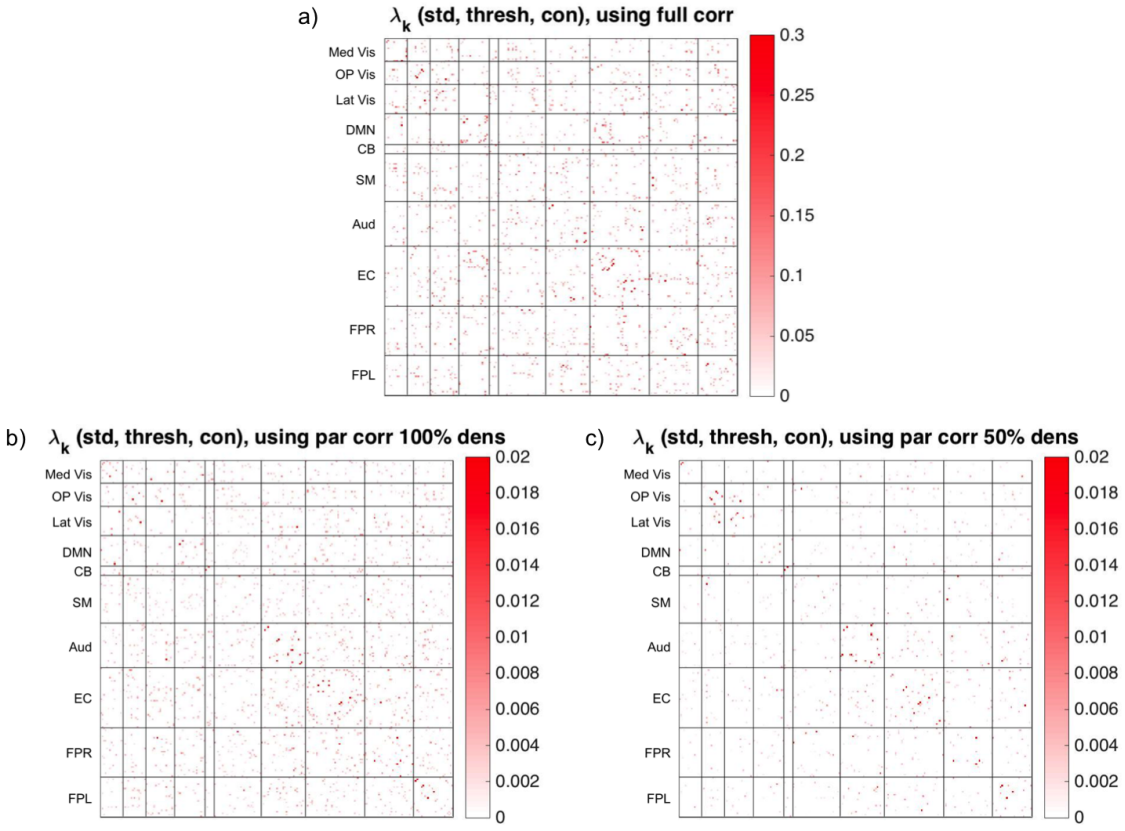
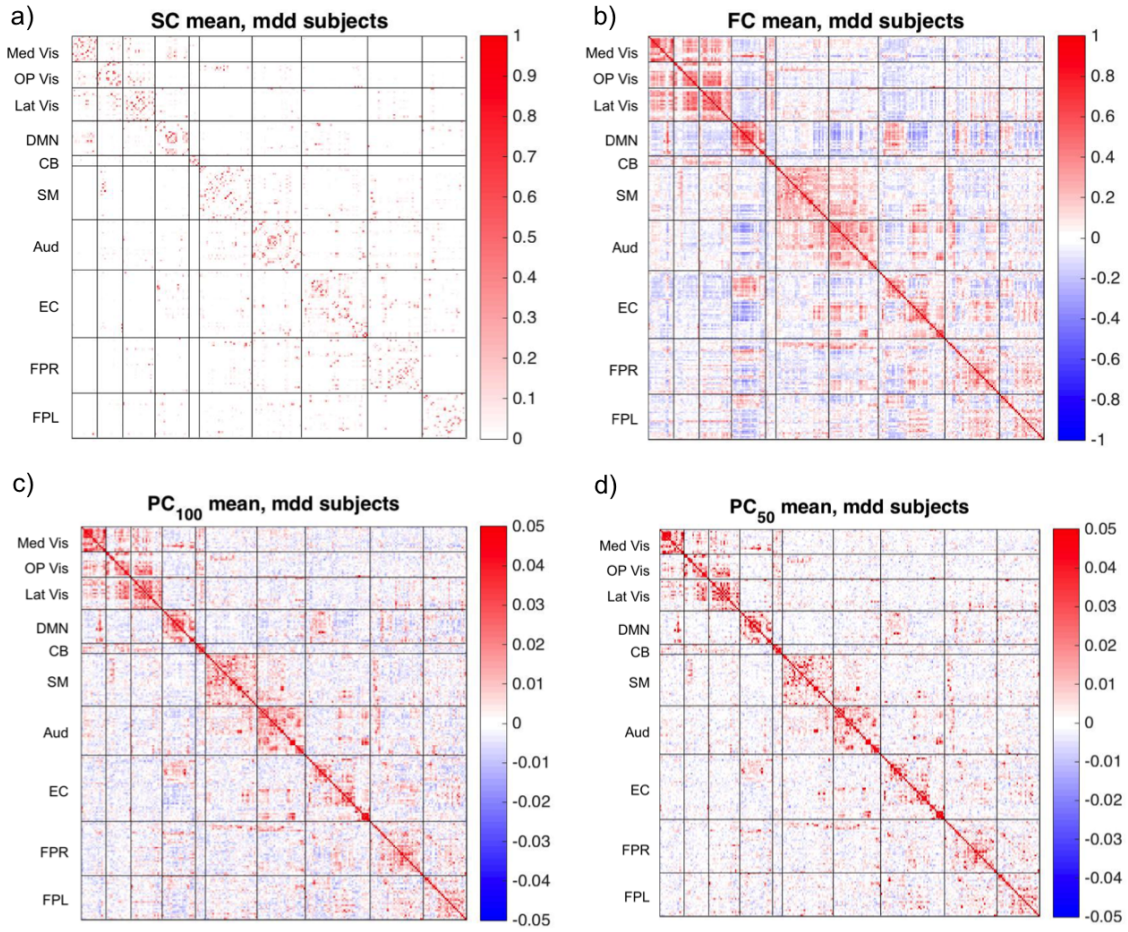


Figure B.2: Mean connectivity matrices, for MDD subjects



Appendix C

Appendix for Chapter 4 (Topic 3)

Figure C.1: Gaussian copula, under different association levels

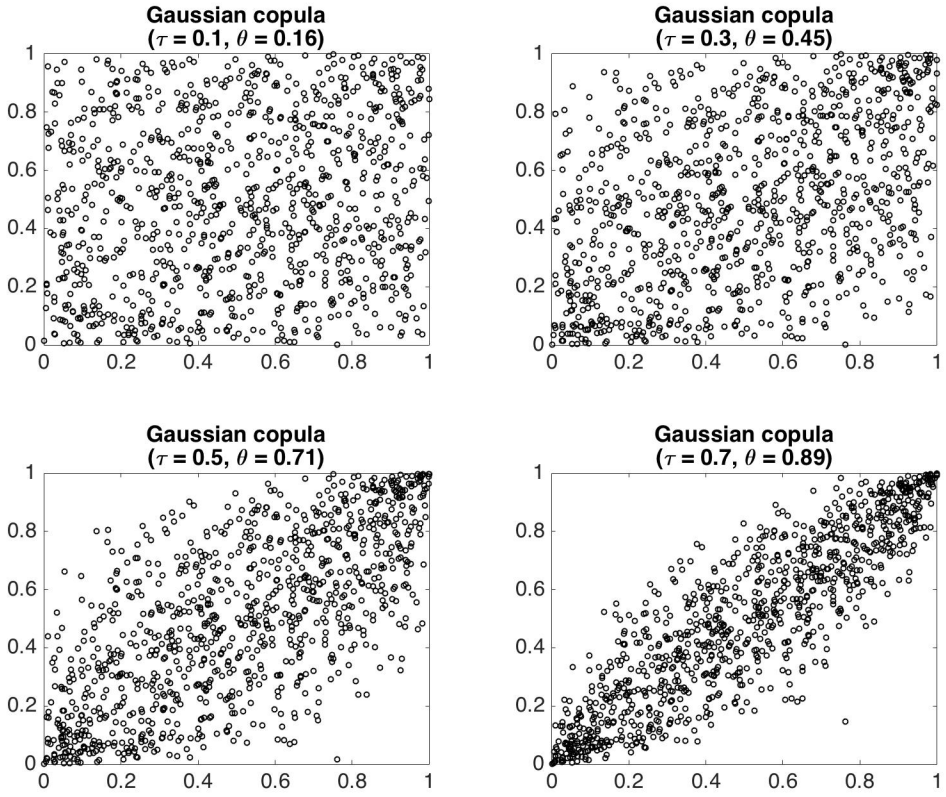


Figure C.2: Student's t copula, under different association levels

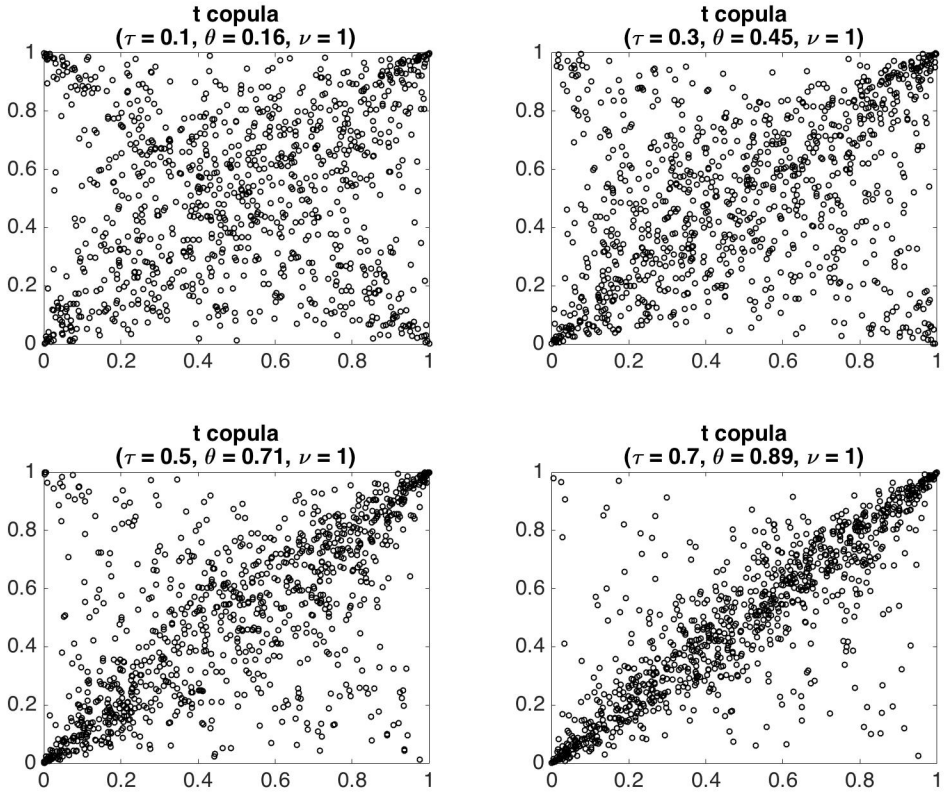


Figure C.3: Clayton copula, under different association levels

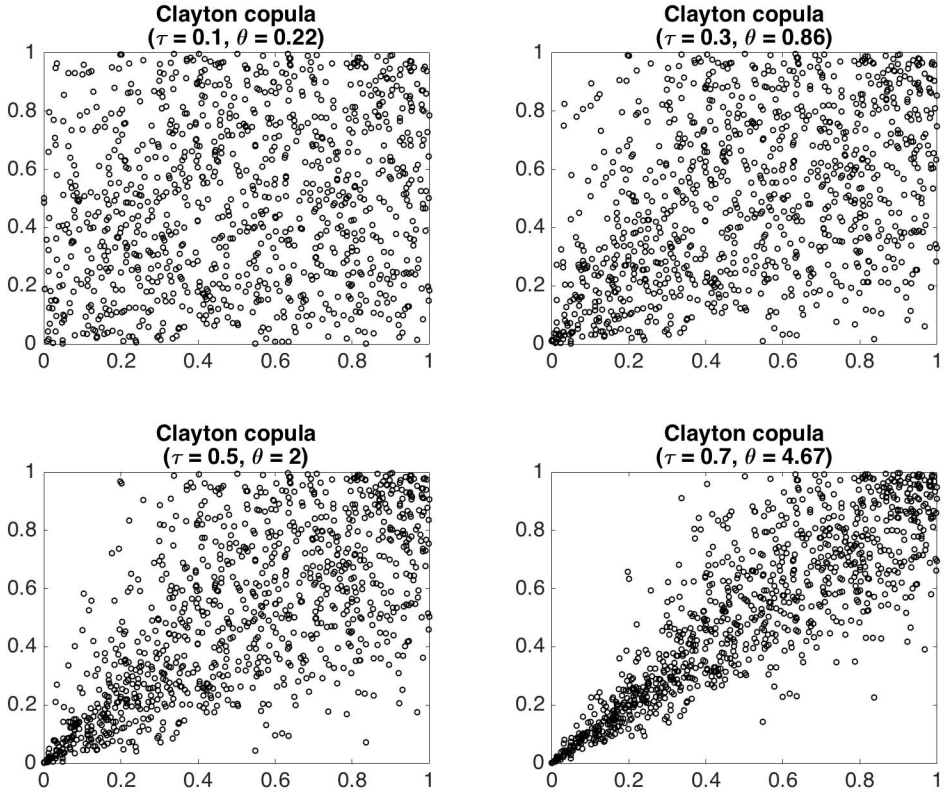


Figure C.4: Gumbel copula, under different association levels

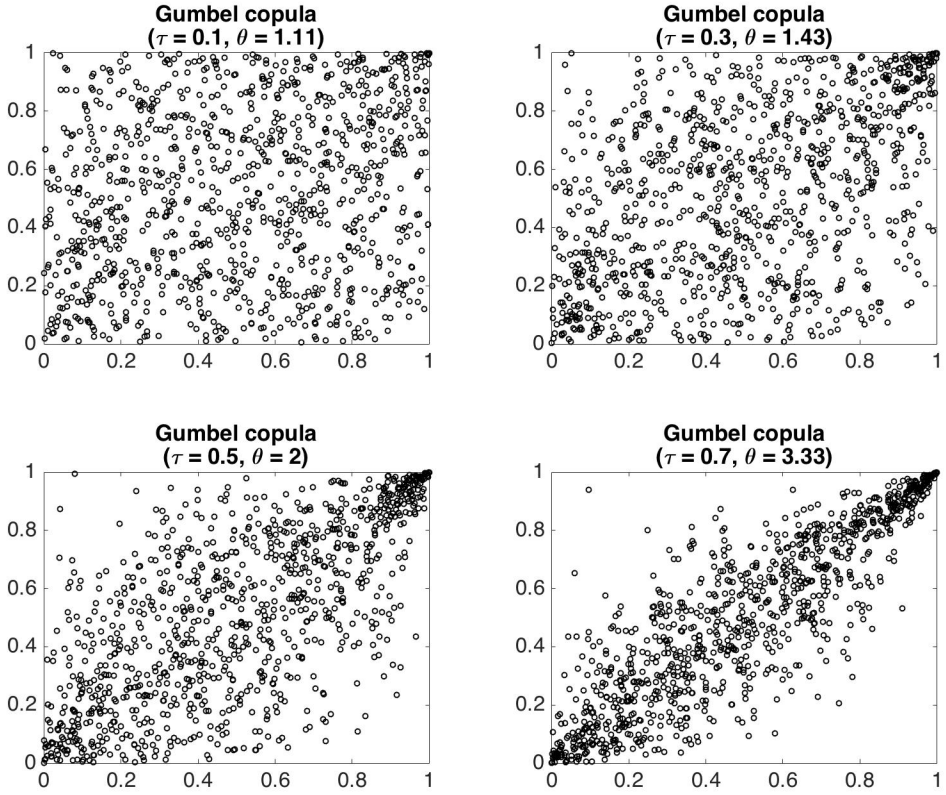


Figure C.5: Frank copula, under different association levels

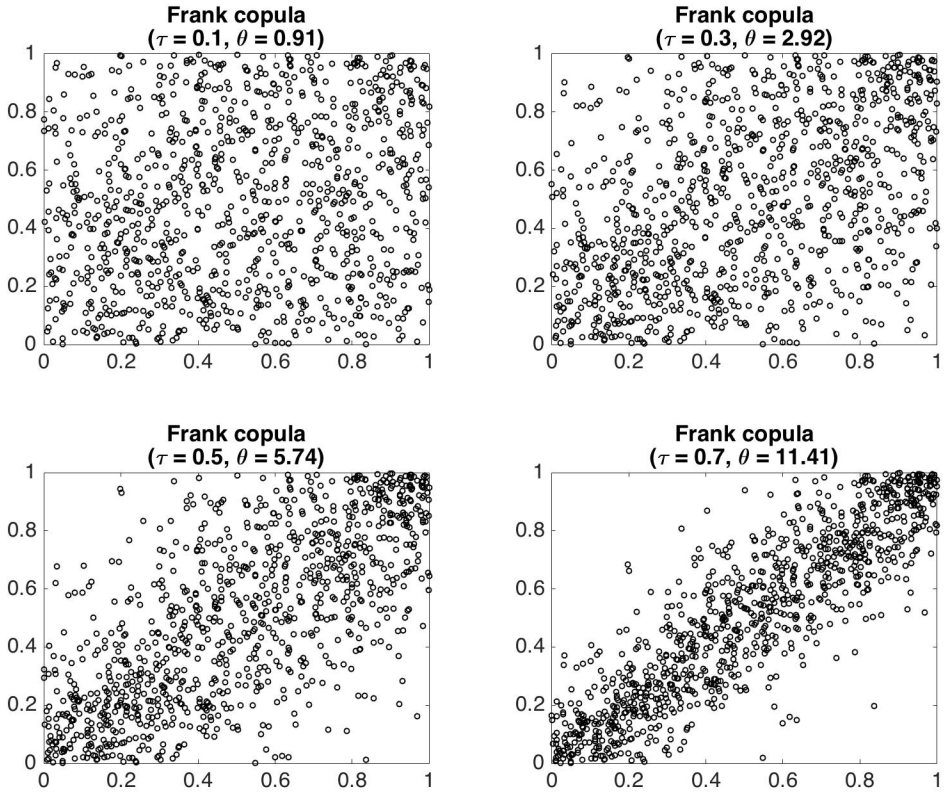


Figure C.6: Edgewise FC-SC association, measured by the Gaussian copula

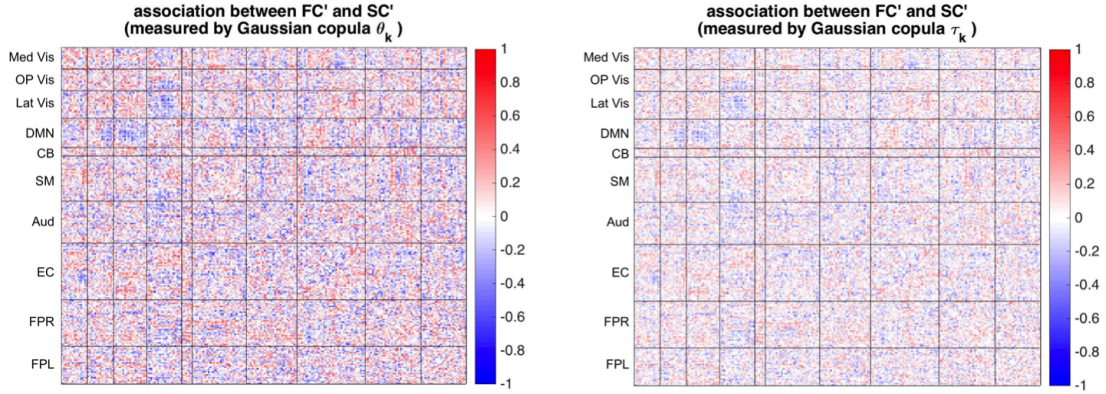


Figure C.7: Edgewise FC-SC association, measured by the t copula

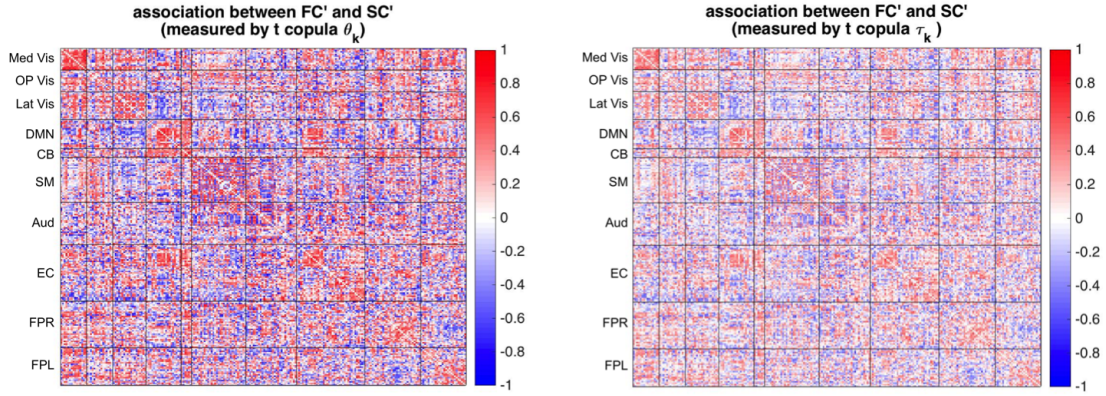
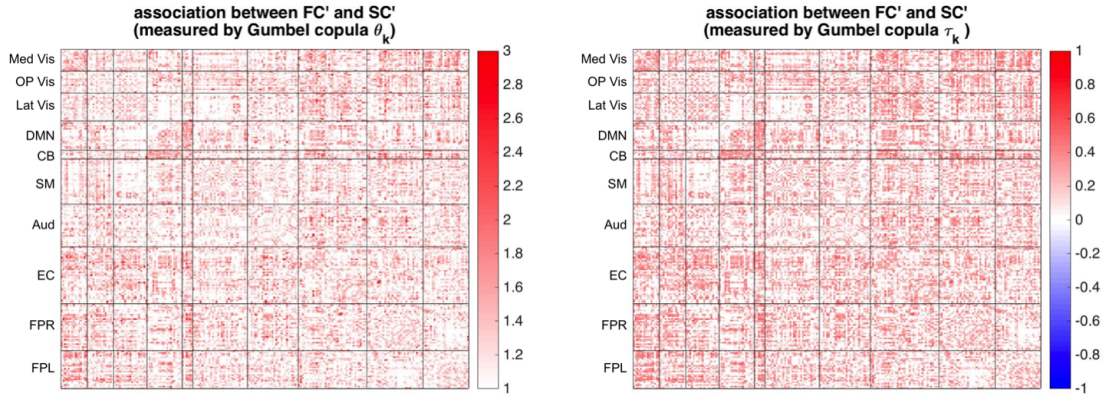


Figure C.8: Edgewise FC-SC association, measured by the Gumbel copula



Bibliography

Almeida, K. J., Brucki, S. M. D., Duarte, M. I. S., Pasqualucci, C. A. G., Rosemberg, S. and Nitrini, R. (2012), 'Basal ganglia lesions in subacute sclerosing panencephalitis', *Dement Neuropsychol* **6**(4), 286–289.

Anand, A., Li, Y., Wang, Y., Lowe, M. J. and Dzemidzic, M. (2009), 'Resting state corticolimbic connectivity abnormalities in unmedicated bipolar disorder and unipolar depression', *Psychiatry Research - Neuroimaging* **171**(3), 189–198.

Bear, M. F., Connors, B. W. and Paradiso, M. A. (2007), *Neuroscience: Exploring the Brain*, 3rd editio edn, Lippincott Williams& Wilkins, Philadelphia, PA.

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

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

Behrens, T. E. J., Berg, H. J., Jbabdi, S., Rushworth, M. F. S. and Woolrich, M. W. (2007), 'Probabilistic diffusion tractography with multiple fibre orientations: What can we gain?', *Neuroimage* **34**(1), 144–155.

Behrens, T. E. J., Woolrich, M. W., Jenkinson, M., Johansen-Berg, H., Nunes, R. G., Clare, S., Matthews, P. M., Brady, J. M. and Smith, S. M. (2003), 'Characterization

and Propagation of Uncertainty in Diffusion-Weighted MR Imaging', *Magnetic Resonance in Medicine* **50**(5), 1077–1088.

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

Bowman, F. D. (2014), 'Brain Imaging Analysis', *Annual Review of Statistics and Its Application* **1**(1), 61–85.

Bowman, F. D., Guo, Y. and Derado, G. (2007), 'Statistical approaches to functional neuroimaging data', *Neuroimaging Clinics of North America* **17**(4), 441–458.

Bowman, F. D., Zhang, L., Derado, G. and Chen, S. (2012), 'Determining functional connectivity using fMRI data with diffusion-based anatomical weighting', *NeuroImage* .

Bozzali, M., Parker, G. J. M., Serra, L., Embleton, K., Gili, T., Perri, R., Caltagirone, C. and Cercignani, M. (2011), 'Anatomical connectivity mapping: a new tool to assess brain disconnection in Alzheimer's disease', *Neuroimage* **54**(3), 2045–2051.

Brown, J. a., Rudie, J. D., Bandrowski, A., Van Horn, J. D. and Bookheimer, S. Y. (2012), 'The UCLA multimodal connectivity database: a web-based platform for brain connectivity matrix sharing and analysis.', *Frontiers in neuroinformatics* **6**(November), 28.

Buckner, R. L., Andrews-Hanna, J. R. and Schacter, D. L. (2008), 'The brain's default network: anatomy, function, and relevance to disease.', *Annals of the New York Academy of Sciences* **1124**(1), 1–38.

Bullmore, E. and Sporns, O. (2009), 'Complex brain networks: graph theoreti-

cal analysis of structural and functional systems', *Nature Reviews Neuroscience* **10**(3), 186–198.

Calhoun, V. D., Adali, T., Giuliani, N. R., Pekar, J. J., Kiehl, K. A. and Pearlson, G. D. (2006), 'Method for multimodal analysis of independent source differences in schizophrenia: Combining gray matter structural and auditory oddball functional data', *Human Brain Mapping* **27**(1), 47–62.

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

Calhoun, V., Pearlson, G. and Adali, T. (2004), 'Independent component analysis applied to fMRI data: a generative model for validating results', *Journal of VLSI signal processing systems for signal, image and video technology* **37**(2-3), 281–291.

Casella, G. and Berger, R. L. (1990), *Statistical inference*, number 241-245, Duxbury Press Belmont, CA.

Cherubini, U., Luciano, E. and Vecchiato, W. (2013), *Copula Methods in Finance*, Wiley Finance.

Collin, G., Sporns, O., Mandl, R. C. W. and van den Heuvel, M. P. (2014), 'Structural and Functional Aspects Relating to Cost and Benefit of Rich Club Organization in the Human Cerebral Cortex.', *Cerebral Cortex* **24**(9), 2258–2267.

Comon, P. (1994), 'Independent component analysis, a new concept?', *Signal processing* **36**(3), 287–314.

Correa, N. and Li, Y. (2009), 'Fusion of fMRI, sMRI, and EEG data using canonical correlation analysis', *Acoustics, Speech and ...* pp. 385–388.

Damoiseaux, J. S. and Greicius, M. D. (2009), 'Greater than the sum of its parts: a review of studies combining structural connectivity and resting-state functional connectivity', *Brain Structure and Function* **213**(6), 525–533.

Damoiseaux, J. S., Rombouts, S., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M. and Beckmann, C. F. (2006), 'Consistent resting-state networks across healthy subjects', *Proceedings of the National Academy of Sciences* **103**(37), 13848–13853.

De Kwaasteniet, B., Ruhe, E., Caan, M., Rive, M., Olabarriaga, S., Groefsema, M., Heesink, L., Van Wingen, G. and Denys, D. (2013), 'Relation between structural and functional connectivity in major depressive disorder', *Biological Psychiatry* **74**(1), 40–47.

Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977), 'Maximum likelihood from incomplete data via the EM algorithm', *JOURNAL OF THE ROYAL STATISTICAL SOCIETY, SERIES B* **39**(1), 1–38.

Drachman, D. A. (2005), 'Do we have brain to spare?', *Neurology* **64**(12), 2004–2005.

Duann, J.-R., Jung, T.-p., Sejnowski, T. J., Makeig, S. and Sets, A. F. D. (2006), Repeated decompositions reveal the stability of infomax decomposition of fMRI data, *in* 'Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the', IEEE, pp. 5324–5327.

Edelman, R. R. and Warach, S. (1993), 'Magnetic resonance imaging'.

Embrechts, P., Mcneil, A., Straumann, D., Embrechts, P., Mcneil, A. and Straumann, D. (1999), *Correlation and dependence in risk management: properties and pitfalls*, Cambridge University Press.

Evans, A. C., Collins, D. L., Mills, S. R., Brown, E. D., Kelly, R. L. and Peters, T. M. (1993), 3D statistical neuroanatomical models from 305 MRI volumes, *in* 'Nu-

clear Science Symposium and Medical Imaging Conference, 1993., 1993 IEEE Conference Record.', IEEE, pp. 1813–1817.

Filippi, M. and Agosta, F. (2011), 'Structural and functional network connectivity breakdown in Alzheimer's disease studied with magnetic resonance imaging techniques', *Journal of Alzheimer's Disease* **24**(3), 455–474.

Friston, K. J. (1994), 'Functional and effective connectivity in neuroimaging: A synthesis', *Human Brain Mapping* **2**(1-2), 56–78.

Gong, G., He, Y., Concha, L., Lebel, C., Gross, D. W., Evans, A. C. and Beaulieu, C. (2009), 'Mapping anatomical connectivity patterns of human cerebral cortex using in vivo diffusion tensor imaging tractography', *Cerebral Cortex* **19**(3), 524–536.

Greicius, M. (2008), 'Resting-state functional connectivity in neuropsychiatric disorders', *Current opinion in neurology* **21**(4), 424–430.

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

Greicius, M. D., Supekar, K., Menon, V. and Dougherty, R. F. (2009), 'Resting-state functional connectivity reflects structural connectivity in the default mode network', *Cerebral Cortex* **19**(1), 72–78.

Groves, A. R., Beckmann, C. F., Smith, S. M. and Woolrich, M. W. (2011), 'Linked independent component analysis for multimodal data fusion', *NeuroImage* **54**, 2198–2217.

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

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

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

Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J. and Sporns, O. (2008), 'Mapping the structural core of human cerebral cortex', *PLoS biology* **6**(7), e159.

Hagmann, P., Kurant, M., Gigandet, X., Thiran, P., Wedeen, V. J., Meuli, R. and Thiran, J.-P. (2007), 'Mapping human whole-brain structural networks with diffusion MRI', *PloS one* **2**(7), e597.

Hamilton, M. (1960), 'A rating scale for depression.', *Journal of neurology, neurosurgery, and psychiatry* **23**, 56–62.

Herculano-Houzel, S. (2009), 'The human brain in numbers: a linearly scaled-up primate brain.', *Frontiers in human neuroscience* **3**(November), 31.

Himberg, J., Hyvärinen, A. and Esposito, F. (2004), 'Validating the independent components of neuroimaging time series via clustering and visualization', *Neuroimage* **22**(3), 1214–1222.

Honey, C. J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J.-P., Meuli, R. and Hagmann, P. (2009), 'Predicting human resting-state functional connectivity from structural connectivity', *Proceedings of the National Academy of Sciences* **106**(6), 2035–2040.

Horn, A., Ostwald, D., Reisert, M. and Blankenburg, F. (2013), 'The structural-functional connectome and the default mode network of the human brain', *NeuroImage* **102**, 142–151.

Jenison, R. L. and Reale, R. A. (2004), 'The shape of neural dependence.', *Neural computation* **16**(4), 665–72.

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W. and Smith, S. M. (2012), 'Fsl', *NeuroImage* **62**(2), 782–790.

Joe, H. (2015), *Dependence Modeling with Copulas*, CRC press, Boca Raton, FL.

Johansen-Berg, H. and Rushworth, M. F. S. (2009), 'Using diffusion imaging to study human connectional anatomy.', *Annual review of neuroscience* **32**, 75–94.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K. R., Rush, A. J., Walters, E. E. and Wang, P. S. (2003), 'The epidemiology of major depressive disorder', *JAMA: the journal of the American Medical Association* **289**(23), 3095–3105.

Kostova, S. P., Rumchev, K. V., Vlaev, T. and Popova, S. B. (2012), 'Using Copulas to Measure Association between Air Pollution and Respiratory Diseases', *Proceedings of World Academy of Science, Engineering and Technology* **6**(January 2012), 749–754.

Kumar, P. and Shoukri, M. M. (2007), 'Copula based prediction models: an application to an aortic regurgitation study.', *BMC medical research methodology* **7**(February), 21.

Kumar, P. and Shoukri, M. M. (2008), 'Evaluating Aortic Stenosis Using the Archimedean Copula', *Journal of Data Science* **6**(January), 173–187.

Laird, A. R., Eickhoff, S. B., Li, K., Robin, D. A., Glahn, D. C. and Fox, P. T. (2009), 'Investigating the functional heterogeneity of the default mode network using coordinate-based meta-analytic modeling', *The Journal of Neuroscience* **29**(46), 14496–14505.

Laird, A. R., Fox, P. T. M. T., Eickhoff, S. B., Turner, J. A., Ray, K. L., Mckay, D. R., Glahn, D. C., Beckmann, C. F., Smith, S. M., Fox, P. T. M. T., Kimberly, L., Mckay, D. R., Glahn, D. C., Beckmann, C. F., Smith, S. M. and Fox, P. T. M. T. (2011), 'Behavioral Interpretations of Intrinsic Connectivity Networks', *Journal of Cognitive Neuroscience* **23**(12), 4022–4037.

McKeown, M. J., Hansen, L. K. and Sejnowsk, T. J. (2003), 'Independent component analysis of functional MRI: What is signal and what is noise?', *Current Opinion in Neurobiology* **13**, 620–629.

McKeown, M. J., Makeig, S., Brown, G. G., Jung, T.-P., Kindermann, S. S., Bell, A. J. and Sejnowski, T. J. (1998), 'Analysis of fMRI data by blind separation into independent spatial components', *Human Brain Mapping* **6**(3), 160–88.

Meinecke, F., Ziehe, A., Kawanabe, M. and Muller, K.-R. (2002), 'A resampling approach to estimate the stability of one-dimensional or multidimensional independent components', *Biomedical Engineering, IEEE Transactions on* **49**(12), 1514–1525.

Menon, V. (2011), 'Large-scale brain networks and psychopathology: A unifying triple network model', *Trends in Cognitive Sciences* **15**(10), 483–506.

Meyer-Baese, a., Wismueller, A. and Lange, O. (2004), 'Comparison of two exploratory data analysis methods for fMRI: unsupervised clustering versus independent component analysis.', *IEEE transactions on information technology in*

biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society **8**(3), 387–398.

Minasny, B. and McBratney, A. B. (2005), 'The Mat{é}rn function as a general model for soil variograms', *Geoderma* **128**(3), 192–207.

Nelsen, R. B. (2006), *An Introduction to Copulas*, 2nd editio edn, Springer Science+Business Media, Inc., New York, NY.

Northoff, G., Wiebking, C., Feinberg, T. and Panksepp, J. (2011), 'The 'resting-state hypothesis' of major depressive disorder-a translational subcortical-cortical framework for a system disorder.', *Neuroscience and biobehavioral reviews* **35**(9), 1929–45.

Onken, A., Grünewälder, S., Munk, M. H. J. and Obermayer, K. (2009), 'Analyzing short-term noise dependencies of spike-counts in macaque prefrontal cortex using copulas and the flashlight transformation.', *PLoS computational biology* **5**(11), e1000577.

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

Raichle, M. E. (2011), 'The restless brain.', *Brain connectivity* **1**(1), 3–12.

Richiardi, J., Eryilmaz, H., Schwartz, S., Vuilleumier, P. and Van De Ville, D. (2011), 'Decoding brain states from fMRI connectivity graphs.', *NeuroImage* **56**(2), 616–26.

Rubinov, M. and Sporns, O. (2010), 'Complex network measures of brain connectivity: uses and interpretations', *NeuroImage* **52**(3), 1059–69.

Rudie, J. D., Brown, J. a., Beck-Pancer, D., Hernandez, L. M., Dennis, E. L., Thompson, P. M., Bookheimer, S. Y. and Dapretto, M. (2013), 'Altered functional and structural brain network organization in autism', *NeuroImage: Clinical* **2**, 79–94.

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

Schölzel, C. and Friederichs, P. (2008), 'Multivariate non-normally distributed random variables in climate research-introduction to the copula approach', *Nonlinear Processes in Geophysics* **15**(5), 761–772.

Schweizer, B. and Wolff, E. F. (1981), 'On Nonparametric Measures of Dependence for Random Variables', *The Annals of Statistics* **9**(4), 879–885.

Shih, J. H. and Louis, T. A. (1995), 'Inferences on the Association Parameter in Copula Models for Bivariate Survival Data', *Biometrics* **51**(4), 1384–1399.

Silva, R. F., Plis, S. M., Adali, T. and Calhoun, V. D. (2014), 'A statistically motivated framework for simulation of stochastic data fusion models applied to multimodal neuroimaging', *NeuroImage* **102**, 92–117.

Sklar, M. (1959), 'Fonctions de répartition à n dimensions et leurs marges.', *Université Paris* **8**, 229–231.

Skudlarski, P., Jagannathan, K., Calhoun, V. D., Hampson, M., Skudlarska, B. A. and Pearlson, G. (2008), 'Measuring brain connectivity: diffusion tensor imaging validates resting state temporal correlations', *Neuroimage* **43**(3), 554–561.

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., Filippini, N., Watkins, K. E., Toro, R., Laird, A. R. and Beckmann, C. F. (2009), 'Correspondence of the brain's functional architecture during activation and

rest.', *Proceedings of the National Academy of Sciences of the United States of America* **106**(31), 13040–5.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E. and Others (2004), 'Advances in functional and structural MR image analysis and implementation as FSL', *Neuroimage* **23**, S208—-S219.

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

Sporns, O., Tononi, G. and Edelman, G. M. (2000), 'Theoretical neuroanatomy: relating anatomical and functional connectivity in graphs and cortical connection matrices', *Cerebral Cortex* **10**(2), 127–141.

Stanley, M. L., Moussa, M. N., Paolini, B. M., Lyday, R. G., Burdette, J. H. and Laurienti, P. J. (2013), 'Defining nodes in complex brain networks.', *Frontiers in computational neuroscience* **7**(November), 169.

Sui, J., Adali, T., Yu, Q., Chen, J. and Calhoun, V. D. (2012), 'A review of multivariate methods for multimodal fusion of brain imaging data.', *Journal of neuroscience methods* **204**(1), 68–81.

Talairach, J. and Tournoux, P. (1988), 'Co-planar stereotaxic atlas of the human brain', *Theime, Stuttgart, Germany* **270**, 132.

Toosy, A. T., Ciccarelli, O., Parker, G. J. M., Wheeler-Kingshott, C. a. M., Miller, D. H. and Thompson, A. J. (2004), 'Characterizing function-structure relationships in the human visual system with functional MRI and diffusion tensor imaging', *NeuroImage* **21**(4), 1452–1463.

Tsiatis, A. and Davidian, M. (2004), 'Joint modeling of longitudinal and time-to-event data: an overview', *Statistica Sinica* **14**, 809–834.

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

Uludag, K., Dubowitz, D. J. and Buxton, R. B. (2005), *Basic Principles of Functional Mri*, Elsevier, San Diego.

Van Den Heuvel, M. P., Mandl, R. C. W., Kahn, R. S. and Hulshoff Pol, H. E. (2009), 'Functionally linked resting-state networks reflect the underlying structural connectivity architecture of the human brain', *Human Brain Mapping* **30**(10), 3127–3141.

Veer, I. M. (2010), 'Whole brain resting-state analysis reveals decreased functional connectivity in major depression', *Frontiers in Systems Neuroscience* **4**(September), 1–10.

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

Wang, Y., Kang, J., Kemmer, P. B. and Guo, Y. (2016), 'An efficient and reliable statistical method for estimating functional connectivity in large scale brain networks using partial correlation', *Frontiers in Neuroscience* **10**(March).

Xia, M., Wang, J. and He, Y. (2013), 'BrainNet Viewer: A Network Visualization Tool for Human Brain Connectomics', *PLoS ONE* **8**.

Xue, W., Bowman, F. D., Pileggi, A. V. and Mayer, A. R. (2015), 'A multimodal approach for determining brain networks by jointly modeling functional and structural connectivity', *Frontiers in Computational Neuroscience* **9**(February), 1–11.

Zalesky, A., Fornito, A. and Bullmore, E. (2012), 'On the use of correlation as a measure of network connectivity', *NeuroImage* **60**(4), 2096–2106.

Zalesky, A., Fornito, A., Harding, I. H., Cocchi, L., Y??cel, M., Pantelis, C. and Bullmore, E. T. (2010), 'Whole-brain anatomical networks: Does the choice of nodes matter?', *NeuroImage* **50**(3), 970–983.

Zhang, D. and Raichle, M. E. (2010), 'Disease and the brain's dark energy', *Nature Reviews Neurology* **6**(1), 15–28.

Zhu, D., Zhang, T., Jiang, X., Hu, X., Chen, H., Yang, N., Lv, J., Han, J., Guo, L. and Liu, T. (2013), 'Fusing DTI and fMRI data: A survey of methods and applications', *NeuroImage* **102**, 184–191.