ABSTRACT

Title of Dissertation: Statistical Network Analysis of
High-Dimensional Neuroimaging Data
With Complex Topological Structures

Tong Lu
Doctor of Philosophy, 2023

Dissertation Directed by: Professor Shuo Chen
Department of Mathematics

This dissertation contains three projects that collectively tackle statistical challenges in the field of high-dimensional brain connectome data analysis and enhance our understanding of the intricate workings of the human brain. Project 1 proposes a novel network method for detecting brain-disease-related alterations in voxel-pair-level brain functional connectivity with spatial constraints, thus improving spatial specificity and sensitivity. Its effectiveness is validated through extensive simulations and real data applications in nicotine addiction and schizophrenia studies. Project 2 introduces a multivariate multiple imputation method specifically designed for voxel-level neuroimaging data in high dimensions based on Bayesian models and Markov chain Monte Carlo processes. According to both synthetic data and real neurovascular water exchange data extracted from a neuroimaging dataset in a schizophrenia study, our method indicates high imputation accuracy and computational efficiency. Project 3 develops a multi-level network model based on graph combinatorics that captures vector-to-matrix associations between brain structural imaging measures and functional connectomic networks. The validity of the proposed model is justified through extensive simulations and a real structure-function imaging dataset from UK Biobank. These three projects contribute innovative methodologies and insights that advance neuroimaging data analysis, including improvements in spatial specificity, statistical

power, imputation accuracy, and computational efficiency when revealing the brain's complex neurological patterns.

STATISTICAL NETWORK ANALYSIS OF HIGH-DIMENSIONAL NEUROIMAGING DATA WITH COMPLEX TOPOLOGICAL STRUCTURES

by

Tong Lu

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2023

Advisory Committee:
Professor Shuo Chen, Chair/Advisor
Professor Vince Lyzinski
Professor Tianzhou Ma
Professor Paul Smith
Professor Xin He

Preface

This dissertation represents the culmination of a research journey spanning several years in the field of brain imaging data analysis and its implications in unlocking the intricate workings of the human brain. The completion of this research endeavor has been made possible through funding from the National Institutes of Health under Award Numbers 1DP1DA04896801, EB008432, and EB008281. It is with immense pride and gratitude that I present this work to the academic community.

The motivation behind this research stemmed from the analytical challenges posed by the complex and entangled nature of neuroimaging data in high dimensions and the need to advance the statistical methodologies in order to disentangle the complex data and further reveal various pathological and structural association mechanisms within brain functional connectome. Through the course of this dissertation, I embarked on three distinct projects, each aimed at addressing specific statistical challenges and offering solutions to the field of neuroscience. These projects have not only introduced novel statistical methodologies on a theoretical level, but have also shed light on their practical applicability in neuroimaging data analysis.

It is my sincerest hope that this dissertation contributes to the field of neuroscience and serves as a stepping stone for future research in statistical network models and understanding human brain connectome. May it inspire further exploration, spark curiosity, and foster innovation in the scientific community.

Acknowledgments

I owe my heartfelt gratitude to all the people who have made this thesis possible. Their unwavering support and contributions have profoundly shaped my PhD experience into one that I will cherish forever.

First and foremost, I would like to express my sincere gratitude to my advisor, Professor Shuo Chen, for granting me an invaluable opportunity to engage in challenging yet immensely fascinating and meaningful projects on brain connectome data over the past five years. His steadfast dedication, guidance, support, and patience have played a crucial role in making this five-year journey exceptionally rewarding and unforgettable. It has been a pleasure to work with and learn from such an extraordinary individual.

I would also like to thank my committee members, Professor Vince Lyzinski, Professor Tianzhou Ma, Professor Paul Smith, and Professor Xin He for graciously agreeing to serve on my thesis committee. Their willingness to dedicate their individual time to reviewing my manuscript and offering constructive feedback has been truly invaluable.

I owe my deepest thanks to my family - my mother and father, who have always provided unconditional love and support. Their immense encouragement has propelled me forward, even in the face of daunting challenges. I am indebted to them for their belief in my abilities and for enabling me to pursue my undergraduate and Ph.D. studies in the United States. Words cannot express the gratitude I feel towards them. Without them, everything I own today would have

remained a distant dream.

Lastly, I am also grateful to my significant half, Luke, whose presence and support have been a constant source of strength throughout my life and academic journey. I am fortunate to have him by my side. I extend my best wishes to him as he embarks on his own pursuit of a Ph.D. degree.

Thank you all for making this five-year journey a magical one.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACC	Anterior Cingulate Cortex
AI	Anterior Insula
ALFF	Amplitude Of Low-Frequency Fluctuation
BG	Basal Ganglia
BH-FDR	Benjamini–Hochberg FDR
BOLD	Blood-Oxygenation-Level Dependent
BSGP	Bipartite Spectral Graph Partitioning
CT	Cortical Thickness
DMN	Default Mode Network
DTI	Diffusion Tensor Imaging
FA	Fractional Anisotropy
FABIA	Factor Analysis for Bicluster Information Acquisition
FC	Functional Connectivity
FCN	Functional Connectomic Networks
FDR	False Discovery Rate
fMRI	Functional Magnetic Resonance Imagine
FPR	False Positive Rate
FWER	Family Wise Error Rate
HIMA	High-Dimensional Multiple Imputation
ICBM	International Consortium for Brain Mapping
ITL	Information Theoretic Learning
IW	Inverse Wishart
KL	Kullback Leibler
MAP	Maximum a Posterior
MAR	Missing At Random
MCAR	Missing Completely At Random
MCMC	Markov Chain Monte Carlo
MDL	Minimum Description Length
MI	Multiple Imputation
MICE	Multivariate Imputation by Chained Equations
MNAR	Missing Not At Random
MOAT	Multilayer Network Association Method
MRI	Magnetic Resonance Imaging
MVN	Multivariate Normal
NNGP	Nearest Neighbor Gaussian Processes
NP	Nondeterministic Polynomial

PCA	Principal Component Analysis
PET	Positron Emission Tomography
PMA	Penalized Matrix Decomposition
RBN	Region-Level Brain Network
RLA	Region-Level Analysis
ROI	Regions Of Interest
rs-fMRI	Resting-State Functional Magnetic Resonance Imaging
SCCA	Sparse Canonical Correlation Analysis
SCCN	Spatially Constrained and Connected Networks
SI	Structural Imaging
SZ	Schizophrenia
TNR	True Negative Rate
TPR	True Positive Rate
vFC	Voxel-wise Functional Connectivity
VLA	Voxel-Level Analysis
wMAE	Weighted Mean Absolute Error
wMBE	Weighted Mean Bias Error
wMSE	Weighted Mean Square Error

Chapter 1: Introduction

Brain imaging data, with its diverse data structures and applications, opens up a realm of possibilities for unlocking the mysteries of the human brain. By unraveling fundamental brain structures and functions, brain imaging techniques provide researchers with valuable insights into complex neurological processes. Statistical analysis of brain imaging data has continuously driven groundbreaking research (Bullmore and Sporns, 2009; Cao et al., 2014; Fornito et al., 2016; Rubinov and Sporns, 2010; Simpson et al., 2013). As both neuroimaging technology and statistical methodology advances, the future holds even greater potential for understanding the brain and its role in human cognition and behavior. Motivated by this immense potential, this dissertation aims to develop three distinct statistical models to systematically disentangle the intricate workings of the human brain, including identifying pathophysiological sub-community patterns in brain functional connectome, robustly imputing missingness in imaging data for further analysis, and revealing systematic association patterns between brain structure and function. The applications of these models help pave the way for further discoveries in neuroscience, and assist clinical predictions concerning disease diagnosis and treatment selection.

1.1 Background of neuroimaging data

Brain imaging data encompasses information obtained through a range of non-invasive imaging techniques, enabling visualization of the brain's structure, function, and connectivity. Commonly utilized imaging techniques include magnetic resonance imaging (MRI), diffusion tensor imaging (DTI), functional magnetic resonance imaging (fMRI), positron emission tomography (PET), and electroencephalography (EEG). MRI produces high-resolution images of the brain's structure, providing valuable physical information such as size, shape, and cortical thickness. DTI assesses the integrity of white matter microstructure by measuring fractional anisotropy (FA). fMRI records dynamic changes in blood flow within different brain regions, facilitating the measurement of localized neural activity and functional connectivity (FC). PET provides information about brain function and metabolism by measuring the distribution of a radioactive tracer. EEG measures the electrical activity of the brain, allowing researchers to study the timing and synchronization of neural processes. All these diverse imaging modalities play essential roles in understanding the complexities of brain activity and contribute to various fields of neuroscience research. By collecting data from these imaging modalities, researchers can capture different aspects of brain activity and organization.

1.1.1 Common data structures

Neuroimaging data can take on various data structures, with the most common ones being:

a) **Volumetric Data**: Volumetric data characterizes a three-dimensional (3D) representation of the brain's structure. It is commonly acquired through MRI scans and provides detailed information about brain anatomy, allowing researchers to study brain regions, their sizes, and

shapes (Milchenko and Marcus, 2013; Reiss et al., 1995; Verellen et al., 2008).

b) **Structural data**: Structural data is related to volumetric data but refers to a broader category of information that characterizes the anatomical properties and organization of the brain. It includes measures such as cortical thickness, surface area, volume of brain regions, and connectivity patterns. In statistical analysis, structural data are often stored in vectors (Bullmore and Sporns, 2009; Derado et al., 2010; Smith et al., 2004). For example, a vector $\boldsymbol{X} = \{x_k\}_{k=1}^m$ stores a list of m integrity measures on different white matter tracts.

c) **Functional Connectivity Data**: Functional connectivity data, derived from fMRI or EEG, examines the temporal correlation between different brain regions. It provides insights into how brain regions communicate and work together, enabling researchers to understand brain networks and their involvement in various cognitive processes. In statistical analysis, functional connectivity data is often stored in a binary or weighted adjacency matrix $\boldsymbol{Y}_{n\times n}$ (Penny et al., 2011; Wig et al., 2014; Xia and Li, 2017), where each element $\{y_{ij}\}_{1\le i<j\le n}$ characterizes the strength of functional connectivity between brain regions i and j.

d) **Graph Data**: Graph-based data structures, such as $G = (V, E)$, represent brain networks as nodes (brain regions) and edges (connections between regions), where V denotes the node set with a size of, for example, $|V| = n$, and E denotes the edge set with a size of $|E| = \binom{n}{2}$. Graph analysis allows researchers to study the network properties of the brain, including node centrality, community structure, and information flow (Fornito et al., 2016; Loewe et al., 2014; Zalesky et al., 2010). This approach is particularly useful for understanding the complex organization and functioning of brain systems.

1.1.2 Biological significance

Brain imaging data has revolutionized neuroscience research, providing researchers with unprecedented opportunities to explore and elucidate crucial aspects of the brain and its disorders. Brain imaging data typically contributes to research in the following ways:

a) **Studying Brain Function**: Functional brain imaging techniques like fMRI and EEG help researchers investigate brain activity during various tasks, providing insights into cognitive processes, perception, attention, and memory (Cole et al., 2010; Poldrack, 2008).

b) **Mapping Brain Connectivity**: Brain imaging data allows the mapping of structural and functional connections between different brain regions. This helps researchers understand how information is transmitted and processed within the brain, leading to discoveries about functional networks and their roles in different behaviors and disorders (Drevets et al., 2008; Kemmer et al., 2018; Smith et al., 2004).

c) **Unraveling Neurological Disorders**: Brain imaging data aids in studying neurological and psychiatric disorders (Siuly and Zhang, 2016; Tae et al., 2018). By comparing brain images of healthy individuals and patients, researchers can identify structural and functional abnormalities associated with conditions such as Alzheimer's disease, schizophrenia, and depression. This knowledge enhances early detection, treatment, and monitoring of these disorders.

d) **Personalized Medicine and Brain-Computer Interfaces**: Brain imaging data can contribute to personalized medicine by providing individualized information about brain structure and function (Dilsizian and Siegel, 2014; Eckelman et al., 2008; Lambin et al., 2017). It also plays a vital role in developing brain-computer interfaces, allowing direct communication between the brain and external devices.

1.2 Research questions and literature review

1.2.1 Voxel-level and region-level analysis

In this work, we aim to statistically study neuroimaging data from two perspectives: voxel-level and region-level analysis, which are two major categories in this domain.

Voxel-Level Analysis (VLA)

VLA is a widely used technique that enables researchers to investigate brain activity and connectivity at the level of individual voxels, which are three-dimensional pixels that compose an image. The most commonly employed method for VLA is fMRI. fMRI measures changes in blood oxygenation levels as a proxy for neuronal activity, providing researchers with insights into brain function.

One of the key advantages of VLA is its ability to capture fine-grained spatial details. Researchers can identify specific brain regions that exhibit significant activation or deactivation during different cognitive processes or in response to external stimuli. VLA has been instrumental in advancing our understanding of brain function and its relationship to various cognitive and psychological processes.

Region-Level Analysis (RLA)

RLA takes a broader perspective by grouping voxels into anatomically defined brain regions. This approach aims to understand the overall functioning of specific brain regions or networks rather than focusing on individual voxels. RLA is commonly used in both structural and functional

brain imaging studies.

In structural brain imaging, such as MRI, RLA involves segmenting the brain into anatomical regions of interest. By quantifying the volume, shape, or cortical thickness of these regions, researchers can investigate structural differences associated with various neurological conditions or developmental changes. Functional RLA is often performed using resting-state fMRI data, which captures spontaneous brain activity in the absence of explicit tasks. The data is processed to identify functional connectivity patterns between different brain regions. Various techniques, including seed-based correlation analysis, independent component analysis (ICA), and graph theory approaches, are used to map and quantify these functional networks.

RLA provides a macroscopic view of brain organization and interregional communication. It allows researchers to study large-scale brain networks and investigate how these networks contribute to various cognitive processes, such as attention, memory, and decision-making.

1.2.2 Research questions

We are interested in three research questions focusing on neuroimaging data. The first two questions delve into the voxel-level analysis, while the final question pertains to region-level analysis:

1. How to build a model that can, on a voxel-pair-level, identify the functionally altered brain sub-area pairs between two large regions caused by a certain brain disease (e.g., Alzheimer's disease, Parkinson's disease, schizophrenia, etc.)?

2. How to construct a missingness imputation technique that is specifically effective for high-dimensional neuroimaging data, which are often resulted from data acquisition limitations

and susceptibility artifacts?

3. How to develop a model that can reveal the underlying systematic association patterns between brain structure (e.g., cortical thickness) and brain function (e.g., functional connectivity between neurons)?

1.2.3 Current methods

The solutions to questions (1) and (3) typically involve population-level multiple testing and covariance/association analysis in order to reveal hidden association patterns. Traditional multiple testing methods, such as the false-discovery rate (FDR) and family-wise error rate (FWER) control, provide statistical safeguards against the inflation of Type I error rates in situations involving multiple comparisons. However, in many applications, these methods can be conservative, resulting in reduced power and potential false negatives. In recent years, alternative approaches, including permutation-based methods and FDR estimation methods, have emerged as alternatives to traditional multiple testing methods. These newer methods offer additional flexibility and adaptability to various research scenarios, addressing some of the limitations associated with traditional approaches. However, these methods frequently do not apply to multivariate voxel/region pairs, as they are unable to account for brain anatomical restrictions and recover inherent systematic patterns of voxels/regions that are associated with covariates of interest, such as clinical status and brain structural measures.

In recent years, researchers have introduced advanced statistical methods to extract sub-community structures while addressing the need for multiple corrections. For instance, Xia and Li (2017) developed a localized statistical inference approach that takes into account network

properties. Chen et al. (2016) proposed a Bayesian hierarchical model to identify voxel-level connectivity patterns related to clinical covariates, and subsequently used voxel-wise functional connectivity (vFC) patterns to infer region-level connections. These innovative approaches have shown enhanced inference results and localized specificity. However, these methods have limited compatibility in integrating spatial information, thereby presenting a potential area for improvement.

In addition, other advanced statistical methods have been developed to jointly model two sets of neuroimaging features by leveraging techniques such as regularization, low rank, and projection models (Kong et al., 2019; Li et al., 2012b; Wang et al., 2011; Zhu et al., 2014). These methods have been successfully applied in multimodal imaging data analysis, yielding intriguing findings (Ball et al., 2017; Hayden et al., 2006; Wehrle et al., 2020; Zhang et al., 2022). These statistical methods can be broadly classified into two categories. The first category utilizes regularization-based methods (Wang et al., 2020; Zhou and Li, 2014; Zhu et al., 2017) that aim to select a parsimonious set of associations between FC and neuroimaging features. However, a major limitation of these methods is their failure to consider the systematic impact of the features on the functional connectivity network. The second category employs dimensionality reduction strategies, such as principal component analysis (PCA) (Chachlakis et al., 2019; Hotelling, 1933; Jolliffe and Cadima, 2016). These methods project both FCs and neuroimaging features onto a reduced set of principal components, followed by regression analysis. However, as an unsupervised dimension reduction technique, PCA-based analysis often captures fewer relevant principal components, resulting in the omission of underlying true association pairs. Sparse canonical correlation analysis (sCCA) methods, which integrate elements from both categories, have gained popularity (Lin et al., 2013; Uurtio et al., 2019; Witten et al., 2009). However, sCCA methods typically operate as

vector-to-vector association analyses and may overlook systematic vector-to-network association patterns. As a result, a methodological gap persists in effectively modeling vector-to-matrix associations, such as the associations between the neuroimaging feature vector and the FCN matrix-variate outcome, while incorporating latent topological network structures.

Lastly, question (2) involves the problem of missingness in neuroimaging data, which commonly arises in neuroimaging studies due to data acquisition limitations and susceptibility artifacts. Simply omitting missing entries may lead to the exclusion of areas of particular interest and reduce statistical power. The mean/mode imputation is one of the most commonly used imputation method, involving replacing missing values with the mean (for continuous variables) or the mode (for categorical variables) of the observed data. While straightforward to implement, mean/mode imputation assumes that the missing values share the same statistical characteristics as the observed values. To improve upon simple imputation, multiple imputation offers a powerful and flexible approach. MI generates multiple plausible imputations for missing values based on observed data. Each imputation is analyzed separately, and the results are combined using specific rules to obtain valid statistical inferences. MI considers the uncertainty associated with imputed values and provides more reliable estimates compared to single imputation methods. Furthermore, the Expectation-Maximization (EM) algorithm is an iterative procedure used for estimating missing values based on maximum likelihood estimation. It assumes that the data are missing at random (MAR) and iteratively estimates the missing values until convergence. It is worth noting that neuroimaging data is typically stored in high dimensions, which can pose challenges for existing imputation techniques, including the intractability of large matrix sampling and high computational complexity. Therefore, there is a potential for improvement in specifically addressing the imputation of high-dimensional brain imaging data.

1.3 Proposed methods

In this dissertation, I present three statistical methods that address each of the questions raised in the previous section (see Section 1.2.2). Each method is specifically designed to tackle the respective research question, and I provide a brief overview of the motivation, current challenges, proposed solutions, and performance evaluation for each method in the following three subsections.

1.3.1 **S**patially **c**onstrained and **c**onnected **n**etworks (SCCN)

Brain connectome analysis commonly compresses high-resolution brain scans (typically composed of millions of voxels) down to only hundreds of *regions of interest* (ROIs) by averaging within-ROI signals. This huge dimension reduction improves computational speed and the morphological properties of anatomical structures; however, it also comes at the cost of substantial losses in spatial specificity and sensitivity, especially when the signals exhibit high within-ROI heterogeneity. Oftentimes, abnormally expressed *functional connectivity* (FC) between a pair of ROIs caused by a brain disease is primarily driven by only small subsets of voxel pairs within the ROI pair. This article proposes a new network method for detection of voxel-pair-level neural dysconnectivity with spatial constraints. Specifically, focusing on an ROI pair, our model aims to extract dense sub-areas that contain aberrant voxel-pair connections while ensuring that the involved voxels are spatially contiguous. In addition, we develop sub-community-detection algorithms to realize the model, and the consistency of these algorithms is justified. Comprehensive simulation studies demonstrate our method's effectiveness for reducing the false-positive rate while increasing statistical power, detection replicability, and spatial specificity. We apply our approach to reveal: (i) disrupted voxel-wise FC patterns related to nicotine addiction between

the basal ganglia, hippocampus, and insular gyrus from 3269 participants using UK Biobank data; (ii) voxel-wise schizophrenia-altered FC patterns within the salience and temporal-thalamic network from 330 participants in a schizophrenia study. The detected results align with previous medical findings but include improved localized information.

1.3.2 **H**igh-d**i**mensional **m**ultiple imput**a**tion (HIMA)

Neuroimaging data typically contain missing entries due to data acquisition limitations and susceptibility artifacts. Simply omitting missing entries may exclude areas of particular interest and decrease statistical power. Besides, many existing model-based imputation methods suffer from high-dimensional data due to the intractability of large matrix sampling and high computational complexity. This paper proposes a multivariate multiple imputation method, HIMA, which is particularly designed for high-dimensional neuroimaging data. To account for approximately normally distributed brain signals, HIMA employs a joint multivariate normal model and constructs conditional probabilities based on Bayesian models using Markov chain Monte Carlo processes. While the normal mean vector is Gibbs sampled, HIMA samples the normal covariance matrix from the posterior mode (i.e., *maximum a posterior probability*). We justified that the posterior mode has achieved good asymptotic properties. Given high-dimensional imaging data, the relaxed posterior sampling step largely enhances numerical stability and imputation accuracy while reducing computational complexity from $\mathcal{O}(Cp^3)$ to $\mathcal{O}(Cp)$, where C depends on sample size, number of iterations, etc., and p is the variable space dimension. We evaluated HIMA on two imaging datasets (semi-synthetic and real data) and compared it with commonly used imputation methods. The results showed that HIMA is robust against large datasets ($n \ll p$) and it expanded

brain map coverage with improved imputed results (reduced bias and dispersion) and significantly improved computational efficiency (10^3 times faster than the popular multiple imputation model MICE).

1.3.3 **M**ulti-level netw**o**rk associ**a**tion me**t**hod (MOAT)

The goal of our research is to model the association between brain structural imaging (SI) measures and functional connectomic networks (FCN) derived from neuroimaging data. In this analysis, the outcomes are off-diagonal elements of functional connectivity (FC) matrices, while predictors are a multivariate vector of SI variables and nuisance variables. We propose a vector-to-matrix multi-level network model to capture latent association patterns between subsets of SIs and FC sub-networks. The first layer network is a bipartite graph characterizing the association between all SI variables and FC outcomes, where an edge denotes a non-zero FC-SI association. Previous findings show that a large proportion of edges are often located within dense bipartite subgraphs, while other edges are randomly and sparsely distributed in the rest of the graph. The second layer network represents a connectomic graph, where most FC outcomes from the first layer dense subgraphs comprise dense clique subgraphs. This globally sparse and locally dense multi-level network model helps to reveal which FCN sub-networks are systematically influenced by which subsets of SIs. We develop algorithms to identify the underlying multi-level sub-networks and propose a statistical inference framework to test these sub-networks. We perform extensive simulation analysis to benchmark the validity and performance of the proposed method. We further apply our approach to 4242 participants from UK Biobank to evaluate the effects of whole-brain white matter microstructure integrity and cortical thickness on the whole-brain FCN.

1.4 Organization of the Dissertation

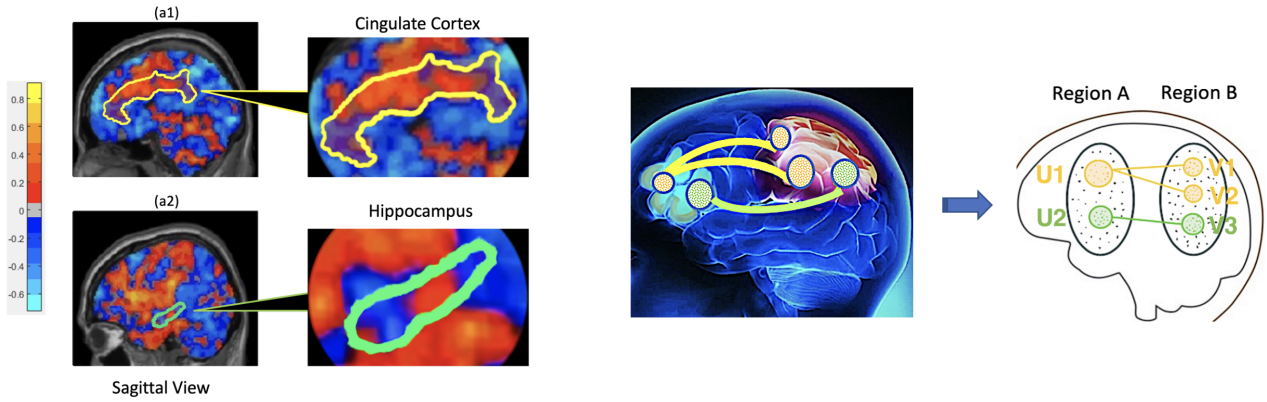
The remaining chapters of this dissertation are organized as follows: Chapter 2 presents the SCCN method, a network-based approach for detecting brain-disease-related alterations in voxel-pair-level brain functional connectivity while incorporating spatial constraints. Chapter 3 presents the HIMA method, a multiple imputation technique specifically designed to address missingness in high-dimensional neuroimaging data. Chapter 4 presents the MOAT method, a multi-level network approach that uncovers the vector-to-matrix associations between brain structural imaging measures and functional connectomic networks.

Chapter 2: Network analysis with spatial-contiguity constraints (SCCN)

2.1 Introduction

Statistical network analysis and graph theory have been fundamental in the study of the intricate neural circuits in human brains (the "human connectome") (Bullmore and Sporns, 2009; Rubinov and Sporns, 2010). A large body of literature has revealed that the human connectome is a well-organized network, and it exhibits graph properties of intelligent networks such as social networks and the Internet (Bahrami et al., 2019; Cao et al., 2014). Built on graph theory, brain network analysis depicts the brain connectome as a graph in which cortical regions are denoted as nodes and the connections between regions are edges. Under this framework, abundant statistical models have been developed to study the associations between complex neural connections and experimental/clinical conditions (e.g., Fornito et al., 2016; Simpson et al., 2013). These models can help to enhance our understanding of the underlying pathophysiological mechanisms of brain diseases (e.g., Alzheimer's disease and Parkinson's disease) and assist clinical predictions concerning disease diagnosis and treatment selection.

In brain network studies, regions of interest (ROIs) are often considered as basic units of analysis, and these are equivalent to nodes/vertices in graph theory. The popularity of region-level brain network (RBN) analysis comes from its high anatomical consistency and computational tractability. When a whole-brain connectome is considered, RBN analysis dramatically reduces



(a) A demonstration of voxel-level variability within ROIs (b) An example of covariate-related sub-area connections

Figure 2.1: *(a) shows the heterogeneity of functional connectivity (FC) among intra-ROI voxels through a seed-to-voxel analysis using insula as a seed ROI. While both the cingulate cortex and hippocampus are well-known ROIs, their interior FC with insula varies substantially. (b) shows a simplified example of covariate-related FC of voxel-pairs located in sub-area pairs* $(U_1, V_1), (U_1, V_2)$, *and* (U_2, V_3) *within a larger ROI pair (Region A, Region B).*

the search dimensions from trillions ($10^6 \times 10^6$) to thousands ($10^2 \times 10^2$). However, RBN analysis relies on the assumption of signal homogeneity among intra-ROI voxels, which is often violated in reality. When significant intra-ROI heterogeneity is present, RBN analysis can lead to several analytical flaws:

1. *Variability negligence*. Simply averaging the time series of voxels within an ROI can lead to voxel-level information variability loss (e.g., Figure 2.1(a));

2. *Spatial specificity loss*. A clinical covariate may alter the ROI-pair connections by disrupting only a small proportion of intra-ROI voxel pairs. In such cases, RBN analysis fails to precisely distinguish the localized alterations;

3. *Power loss*. The averaging process mixes both significant and non-significant voxel-level connections, which often attenuates the effect size and statistical power.

Recently, many brain network studies have shifted focus from RBN analysis to voxel-level network analysis (Loewe et al., 2014; Wu et al., 2013). Traditional multiple testing methods (e.g.,

the false-discovery rate (FDR) and the family-wise error rate (FWER) control) are not applicable to high-dimensional multivariate voxel pairs since they are unable to take into account anatomical restrictions and inherent systematical patterns of disease-associated voxels in ROIs. Some other existing methods may also have limitations, such as not utilizing rich voxel-level information to complement region-level connectivity characterization, or yielding relatively hard-to-interpret results for various reasons (e.g., under-represented neurobiological structures or biases in the seed-selection process). Several advanced statistical methods have been proposed to address these limitations. For example, Xia and Li (2017) provided localized statistical inference by accounting for the network properties. Chen et al. (2016) proposed a Bayesian hierarchical model to identify the voxel-level connectivity patterns associated with clinical covariates and then used the voxel-wise functional connectivity (vFC) patterns to infer region-level connections. These novel approaches yield improved inference results and localized specificity. Nonetheless, they are not directly applicable to our input data of interest (i.e., an $m \times n$ "bi-cluster" rather than an $n \times n$ adjacency matrix), and they do not regulate involved voxels to be spatially contiguous. Unlike RBN analysis, spatial contiguity is crucial for vFC analysis because: (i) it preserves anatomical homogeneity, and it hence preserves the interpretability of the vFC results (Thirion et al., 2006); (ii) it better controls the FDR and FWER since phenotype-related vFC is often intrinsically linked with the topological structure of the brain connectome (Fan et al., 2012).

In this study, our goal was to identify altered vFC patterns between spatially contiguous sub-area pairs from a larger region pair. More specifically, given a region pair of interest, we sought to extract interior sub-area pairs that could maximally cover spatially adjacent covariate-related vFC with well-controlled FDR and FWER values (e.g., Figure 2.1(b)). Our sub-area *extraction* approach is fundamentally distinct from other commonly used brain *parcellation*

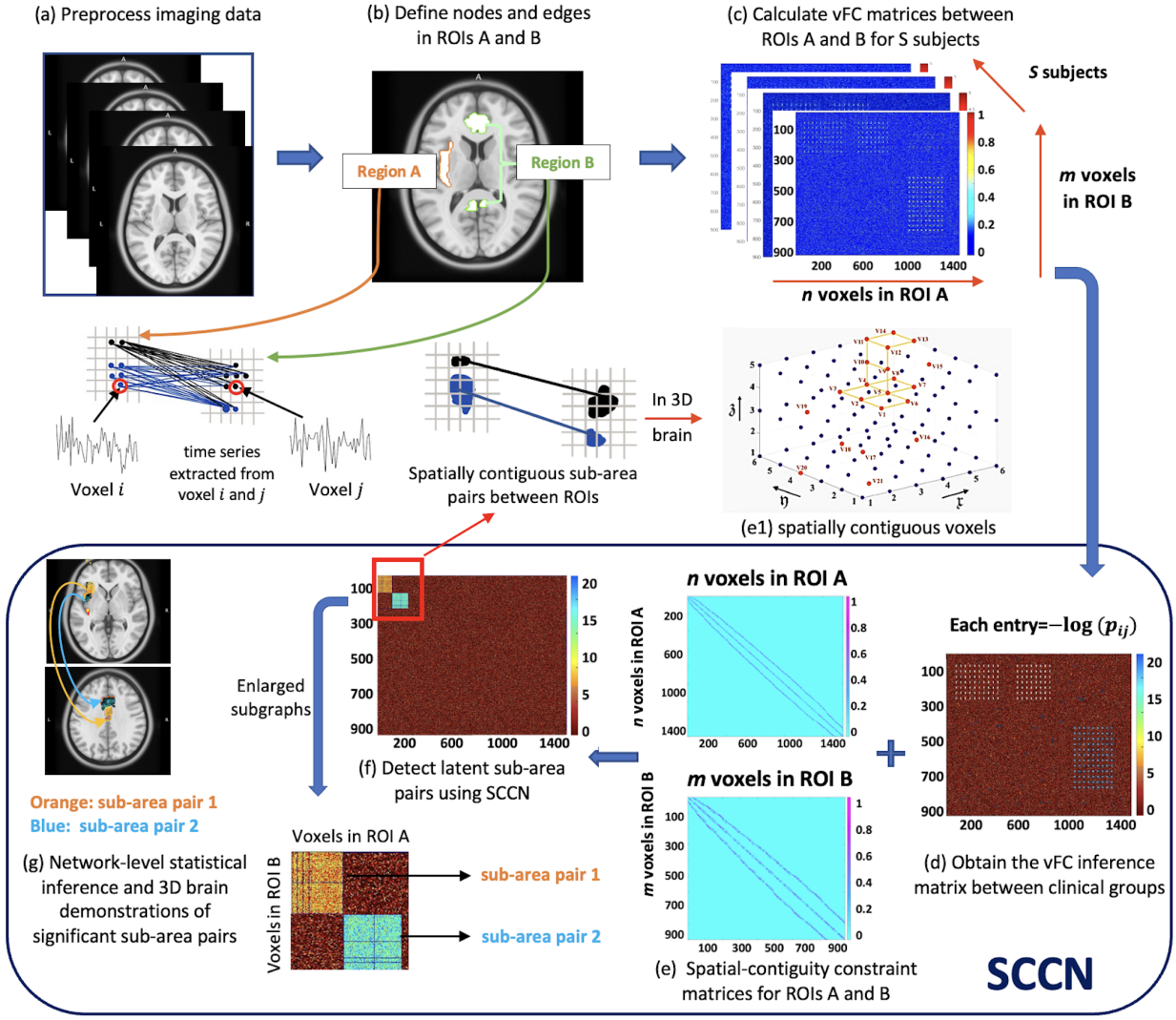


Figure 2.2: *SCCN pipeline.(a) Preprocess the fMRI data and transform it into a standard brain template. (b) Define voxels in ROIs as nodes and bonds between voxels as edges. Extract the time series of brain signals from each voxel. (c) Calculate the connectivity matrix between voxels from regions A and B for each subject. (d) Calculate the connectivity inference matrix, where each element is a test statistic per edge between clinical groups. A hotter point in the heatmap suggests a larger between-group difference. (e) Construct the spatial-contiguity constraint matrices for ROIs A and B (see detailed matrix construction in Section 2.1). In (e1), each dot represents a voxel in 3D coordinates, where red dots represent positive voxels. Voxels connected by yellow lines form a spatially contiguous area. (f) Detect the disease-related connections contained in sub-area pairs based on (d) and (e) jointly. (f) is obtained by re-ordering the nodes in (d), with the densely altered sub-networks pushed to the top (i.e., (d) and (f) are isomorphic graphs). (g) Conduct the proposed MDL-based network-level statistical inference. The sub-area pairs that pass the statistical tests are highlighted.*

methods such as anatomy-based and data-driven approaches (e.g., gradient- or similarity-based mappings) (Craddock et al., 2012; Wig et al., 2014); these parcellation methods seek to segment an ROI into different sub-regions, and every single voxel is assigned to a corresponding sub-region. In contrast to parcellation methods in which every voxel is processed, our sub-area extraction approach only selects subsets of voxels that are covariate-related and are constrained in spatially contiguous spaces. All other non-selected voxels are considered to be covariate-indifferent. Sub-area extraction is more suited to our study because: (i) it is likely that the covariate-related differences across clinical groups may gather in the vFC between a sub-area in Region A and an intersection of multiple sub-areas grouped by the existing parcellation methods in Region B; (ii) it is often found that only a small proportion of voxels in regions A and B are disrupted, and thus a comprehensive parcellation across the entire ROI is not necessary (Cao et al., 2014).

To achieve the desired sub-area extraction and address the limitations discussed above, we propose a new statistical network framework to extract **S**patially **C**onstrained and **C**onnected **N**etworks, hereafter referred to as SCCN. SCCN is a two-step method (Figure 2.2) focusing between a pair of ROIs, say A and B, that are believed to contain aberrant functional connections caused by a brain disease. In **step 1**, SCCN extracts spatially coherent sub-area pairs that maximally contain disease-altered vFC between regions A and B. In **step 2**, we formally test each extracted sub-area pair to determine whether it is significantly covariate-associated with multiple testing controls. If no sub-area pairs are found to be significant, we then consider the region-pair connectivity as covariate-unrelated. If significant results are seen, the association between the covariate of interest and the ROI-pair connections can be traced down to smaller but much more precise sub-areas consisting of extracted voxels. These vFC results may provide insights into understanding the latent neurophysiological mechanisms of diseases.

In this chapter, we show that SCCN provides a consistent estimate for the true community structure in the sense that the error of edge assignments is negligible in large region pairs. We empirically evaluate the performance of SCCN through extensive simulation studies. The results show that SCCN achieves satisfactory performance in increasing statistical power and spatial specificity while controlling the false-positive rate. Notably, SCCN is easily scalable to both small and large datasets. Besides, we apply SCCN to two real data examples: a nicotine-addiction research study using UK Biobank[1] data with 3269 participants, and a schizophrenia research study with 330 participants. Through these applications, we systematically investigate disease-related sub-network structures using SCCN with rigorously controlled FDR.

2.2 Methods

2.2.1 Background

2.2.1.1 Data structure

Given two ROIs, A and B, each consisting of n and m voxels, respectively, the vFC association patterns can be represented by a general $(n+m)\times(n+m)$ outcome matrix. Specifically, the $(n+m)\times(n+m)$ connectivity matrix can be decomposed into three sub-matrices: $n\times n$, $m\times m$, and $n\times m$, which encompass within-A, within-B, and between-region connections information. Herein, we focus on presenting the new methodology for vFC analysis between ROI A and B (i.e., $n\times m$ connectivity matrix), which is motivated by the growing interest in clinical investigations aimed at exploring neuropsychiatric disorder-related inter-regional vFC changes

[1]UK Biobank is a large-scale biomedical database and research resource containing in-depth genetic and health information from half a million UK participants.

(Agosta et al., 2013; Rogers et al., 2007; Wu et al., 2011). Due to space limitation, we provide the statistical framework for within-region vFC analysis (i.e., the $n \times n$ and $m \times m$ connectivity matrices) along with the additional simulations and real data applications in Appendix 2B.

For a subject $s \in [S] := \{1, \ldots, S\}$, let $\mathbf{Z}^{A,s}_{n\times T}$ and $\mathbf{Z}^{B,s}_{m\times T}$ represent the matrices of voxel-level blood-oxygenation-level dependent (BOLD) signals at T different time points for ROIs A and B. The outcome variables are the functional connectivity measures quantified by similarity matrices between the time series of voxels in A and in B. For example, Y^s_{ij}, the connectivity strength between voxel i in A and voxel j in B, can be computed by $Y^s_{ij} = f(Z^{A,s}_{i\cdot}, Z^{B,s}_{j\cdot})$, where $Z^{A,s}_{i\cdot}$ and $Z^{B,s}_{j\cdot}$ are the BOLD time series for voxels i and j, and f is a similarity metric (e.g., Fisher's z-transformed Pearson correlation). Collecting all Y^s_{ij} for each voxel pair $(i, j) \in [n] \times [m]$ gives an inter-region connectivity matrix $\mathbf{Y}^s_{n\times m}$. Additionally, a covariate vector $\mathbf{X}^s_{1\times p}$ is observed for each subject s, and this contains demographic and clinical information.

Our goal is to identify clinical/behavioral-related functional connectivity (FC) patterns at the voxel level. This is because voxel-level findings can reveal altered FC with improved statistical power and enhanced spatial specificity and resolution. To achieve this, multivariate statistical inference is required for the $n \times m$ vFC outcomes (usually in high dimension, e.g., millions) with spatial constraints. We first test the associations between each outcome Y^s_{ij} and a regressor of primary interest $x^s_1 \in \mathbf{X}^s$ (clinical status in our application, e.g., patient or control):

$$\mathbb{E}(Y^s_{ij}|\mathbf{X}^s) = \alpha_0 + x^s_1\beta_{ij} + \mathbf{X}^s_{1\times(p-1)}\boldsymbol{\alpha},$$

where β_{ij} is the coefficient of x^s_1 and $\boldsymbol{\alpha}$ is a coefficient vector for the remaining covariates $\mathbf{X}^s_{1\times(p-1)}$ (e.g., age, ethnicity, etc). We denote $\boldsymbol{\beta} := \{\beta_{ij}\}_{i\in[n],j\in[m]}$ and aim to systematically extract vFC

whose $\boldsymbol{\beta} \neq 0$ with high accuracy. We further summarize the significance levels of $\boldsymbol{\beta}$ by a *connectivity inference matrix* $\mathbf{W}_{n\times m}$. Each entry of $\mathbf{W}_{n\times m}$ is computed by $\mathbf{W}_{ij} = -\log p_{ij}$, where p_{ij} is the p-value for β_{ij}. In neuroimaging statistics, the selection of $\boldsymbol{\beta} \neq 0$ is not only determined by the level of statistical significance but also by spatial constraints. In addition to these two factors, $\boldsymbol{\beta}$ is also intrinsically linked with an underlying $n \times m$ bipartite graph between ROIs A and B. Therefore, we will require both graphic and spatial information to assist in identifying vFC whose $\boldsymbol{\beta} \neq 0$. We present the detailed graphic and spatial constructions as follows.

2.2.1.2 Graph representation

To decipher the complex voxel-pair connectome, we consider a bipartite graph structure $G = \{U, V\}$ underlying the inference matrix $\mathbf{W}_{n\times m}$. The node sets U and V represent voxels in ROIs A and B, respectively, where $|U| = n$ and $|V| = m$. We assume that, after spatial normalization and registration of the fMRI data, all subjects share a common set of nodes, namely, $(U^s, V^s) \equiv (U, V), \forall s \in [S]$.

2.2.1.3 Spatial contiguity

Each node in our dataset corresponds to a voxel at a certain spatial position in 3D brain imaging (e.g., Figure 2.2(e1)). When we map each detected subgroup of voxels back to the 3D brain space, we desire these voxels to emerge as a spatially adjacent cluster (i.e., connected components). Such anticipation, translated into formal language, is referred to as *spatial contiguity*. Specifically, we define an "infrastructure graph" $\mathcal{S}_A$ between all nodes within ROI A to accommodate spatial contiguity. Each entry $S_{ii'}$ in $\mathcal{S}_A$ is a spatial-adjacency indicator variable between voxels i and

i' in ROI A, where $S_{ii'} = 1$ if $d_{ii'} \leq \varepsilon$, and $S_{ii'} = 0$ otherwise ($d_{ii'}$ is the Euclidean distance between voxels i and i'). For example, in a 3D grid space, when ϵ is set to be $\sqrt{3}$, a centroid voxel i in a cube will have 26 surrounding voxels i' such that $S_{ii'} = 1$. We define and interpret $\mathcal{S}_B$ for nodes within ROI B similarly. $\mathcal{S}_A$ and $\mathcal{S}_B$ will be used to prescribe the spatial-contiguity constraints when implementing SCCN. We provide more rigorous mathematical definitions of spatial contiguity, $\mathcal{S}_A$, and $\mathcal{S}_B$ in Appendix 2A.1.

We propose the SCCN model to systematically select vFC of $\beta_{ij} \neq 0$ by jointly considering the information of voxel-pair-level statistical significance, underlying graph structures, and spatial constraints. We integrate these into a weighted graph $\mathcal{G} = \{W, \mathcal{S}_A, \mathcal{S}_B\}$ as the input of our method.

2.2.2 Detecting densely altered sub-area pairs from an ROI pair

2.2.2.1 Spatial-contiguity-constrained objective function

The node set U corresponding to voxels in ROI A can reportedly be partitioned into mutually non-overlapping sub-areas $\{U_c\}$, denoted by $U = \bigoplus_{c=1}^{C} U_c$ (Eickhoff et al., 2015). Similarly, we have $V = \bigoplus_{d=1}^{D} V_d$ for ROI B. In this chapter, we aim to extract sub-area pairs $\{(U_c, V_d)\}$ that dominantly contain disease-related voxel pairs, and we call these "densely altered" sub-area pairs. Formally, a sub-area pair (U_c, V_d) is considered densely altered if $\sum_{(i,j)\in(U_c,V_d)} \frac{I(\beta_{ij}\neq 0)}{|U_c|\,|V_d|} \gg \sum_{(i,j)\in(U_c',V_d')} \frac{I(\beta_{ij}\neq 0)}{|U_c'||V_d'|}$, where U_c' and V_d' are the complements of node sets U_c and V_d. We are therefore inspired to devise a regularized objective function to generate a checkerboard-like network structure underlying the connectivity inference matrix $\mathbf{W}$. This network structure reshuffles $\mathbf{W}$ and reveals densely altered $\{(U_c, V_d)\}$ pairs from (U, V). In addition, we impose spatial contiguity on U_c and V_d to improve biological interpretability and prohibit isolated false

positive edges. Finally, the objective function is formulated as follows:

$$\underset{\substack{C,\ D,\ U=\bigoplus_{c=1}^{C} U_c, V=\bigoplus_{d=1}^{D} V_d \\ (U_c, V_d \text{ subject to spatial contiguity})}}{\operatorname{argmax}} \int \sum_{c=1}^{C}\sum_{d=1}^{D} \left\{ \log \frac{\sum_{i\in U_c, j\in V_d} W_{ij}\cdot I(W_{ij}>r)}{|U_c||V_d|} + \lambda \log(|U_c||V_d|) \right\} g(r)dr, \tag{2.1}$$

where $\lambda \in [0,1]$ is a tuning parameter, r is a threshold below which there is no disease-related effect on W_{ij}, and $g(r)$ is the distribution function for r. Both $g(r)$ and λ can be chosen by prior knowledge or by a data-driven method proposed in Section 2.2.2.2.

The tuning parameter λ falls in the range $[0,1]$: when $\lambda = 0$, maximizing (2.1) is equivalent to maximizing $f_1 = \frac{\sum_{i\in U_c, j\in V_d} W_{ij}\cdot I(W_{ij}>r)}{|U_c||V_d|}$, which is a popular definition for connection density; when $\lambda = 1$, maximizing (2.1) is simply maximizing $f_2 = \sum_{i\in U_c, j\in V_d} W_{ij}\cdot I(W_{ij}>r)$, which quantifies the magnitude of significant voxel pairs contained by the sub-area pair (U_c, V_d). Direct optimization of the connection density f_1 tends to detect a dense subgraph with a minuscule size, while the optimization of f_2 can trigger an oversized subgraph. Theorem 2.2 shows that function (2.1) provides a consistent estimate for the targeted topological structure (collections of edge-induced sub-area pairs) in the sense that the error of edge assignments is negligible in large region pairs. Extensive simulation studies also show that function (2.1) performs well in balancing the size and density when detecting subgraphs.

2.2.2.2 Optimization of objective function (2.1) for given $g(r)$ and λ

In this section, we focus on optimizing function (2.1) for a given configuration of $g(r)$ and λ, which are the density function for the threshold r and the tuning parameter in (2.1). We will then

discuss how to determine $g(r)$ and λ in the next section. Unfortunately, even with a given $g(r)$ and λ, direct optimization of (2.1) is still an NP-hard problem. Therefore, traditional optimization methods, such as gradient descent, cannot be used due to the non-convexity of the problem. Here, we present an alternative strategy for optimizing (2.1). The essential idea is that we integrate $\mathbf{W}$ with the spatial-contiguity constraints and then estimate the targeted community structure using modified spectral clustering algorithms via iterative procedures. As presented earlier, the targeted network structure is $\{U_c, V_d\}$ partitioned from (U, V) (i.e., the collection of edge-induced sub-area pairs, or in other words, the voxel memberships of U_c and V_d), where $U = \bigoplus_{c=1}^{C} U_c$ and $V = \bigoplus_{d=1}^{D} V_d$.

According to the spectral clustering algorithm, applying singular value decomposition to the Laplacian matrix of $\mathbf{W} = \mathbb{U}\mathbf{\Sigma}\mathbb{V}^\top$ and then clustering $\mathbb{U}$ and $\mathbb{V}$ will give partitions of regions A and B, respectively. Now, since $\mathbb{V}$ is the eigenvectors of $\mathbf{W}^\top\mathbf{W}$, spectral clustering on the Laplacian matrix of $\mathbf{W}^\top\mathbf{W}$ will simply give the partitions of Region B. Similarly, spectral clustering on the Laplacian matrix of $\mathbf{W}\mathbf{W}^\top$ will provide the partitions of Region A. Therefore, our community-detection algorithm can be conducted based on $\mathbf{W}\mathbf{W}^\top$ and $\mathbf{W}^\top\mathbf{W}$. Next, to incorporate the spatial-contiguity constraints into the optimization, we make use of the two within-region "infrastructure graphs" $\mathcal{S}_A$ and $\mathcal{S}_B$ introduced earlier. Specifically, we define

$$\mathbf{W_A} = \mathbf{W}\mathbf{W}^\top \odot \mathcal{S}_A \quad \text{and} \quad \mathbf{W_B} = \mathbf{W}^\top\mathbf{W} \odot \mathcal{S}_B, \tag{2.2}$$

where $\odot$ is an element-wise product. As pointed out by Kamvar et al. (2003) and Craddock et al. (2012), $\mathcal{S}_A$ and $\mathcal{S}_B$ force the similarity between all pairs of non-adjacent voxels to zero, which breaks edges between isolated voxels in the graph. Based on this, the n by n matrix $\mathbf{W}_A(ii')$

(where i and i' are two voxels in A) is greater if the voxels in A are spatially adjacent and have a similar profile linking to voxels in Region B. The spatial-contiguity constraints enable our method to produce results that better honor the neurobiological background regarding the coherence of neighboring neuron populations (Thirion et al., 2006).

We can now fit a stochastic block model to $\mathbf{W}_A$ (and another to $\mathbf{W}_B$) using the spectral clustering algorithm and then perform grid search for the optimizer of function (2.1). We further examine whether the estimated $\hat{U}_c$ and $\hat{V}_d$ values satisfy the spatial-contiguity constraints, while empirically we find that the constraints are typically satisfied. There is thus no need to perform any further modification step for the constraints. We formally present our clustering procedure in Algorithm 1.

Algorithm 1 Optimization of objective function (2.1) with given λ

procedure ALGORITHM (Input: λ and $\mathcal{G} = \{\mathbf{W}, \mathcal{S}_{\mathcal{A}}, \mathcal{S}_{\mathcal{B}}\}$)
  function *SCCN.partition* (λ, $\mathcal{G}$)
  for $C = 1, 2, \ldots, |U|$ **do**
    Ratio-cut spectral clustering $\mathbf{W}_A$ into C networks: $U = \bigoplus_{c=1}^{C} U_c$
    for $D = 1, 2, \ldots, |V|$ **do**
      Ratio-cut spectral clustering $\mathbf{W}_B$ into D networks: $V = \bigoplus_{d=1}^{D} V_d$
      Substitute network sets U and V into objective function (2.1), and obtain the output values
    end for
  end for
  return C, D, $U = \bigoplus_{c=1}^{C} U_c$ and $V = \bigoplus_{d=1}^{D} V_d$ that yield the maximum output value
  end function
end procedure

Consistency for subgraph detection. In Lemma 2.1, we first establish that, given true sub-areas numbers C^* and D^*, the solution to optimize the objective function (2.1) provides a consistent estimate for the topological structure of the target community ($\{U_c, V_d\}$) (the collection of edge-induced sub-area pairs) in the sense that false-positive edge assignments are negligible

in very large bipartite graphs $G = (U, V)$, $|U| \to \infty$, and $|V| \to \infty$. Then, we establish the convergence of Algorithm 1 to optimize function (2.1) based on Theorems 2.2 and 2.3. In Theorem 2.2, we prove that our algorithm can provide a consistent estimate of the number of sub-areas, C and D. In Theorem 2.3, we prove that the implementation of Algorithm 1 converges to the optimal solution of the objective function (2.1).

To present the theoretical results, we consider the following settings. Let $\{e_{ij}^1\}$ and $\{e_{ij}^0\}$ be the sets of positive (e.g., disease-related) and negative edges, respectively. For a weighted matrix $\mathbf{W}$, we assume that $w_{ij}|e_{ij}^1 \overset{\text{iid}}{\sim} f_1$ and $w_{ij}|e_{ij}^0 \overset{\text{iid}}{\sim} f_0$, where f_1 and f_0 are two probability density functions with means and variances (μ_1, σ_1^2) and (μ_0, σ_0^2), respectively. In addition, let $\mathcal{M}^*$ be the true membership of edges (the community index of edges falling in sub-area pair (U_c, V_d)). Furthermore, let $\hat{\mathcal{M}}_{(\hat{C},\hat{D})}$ be the membership estimated by function (2.1) with $\hat{C}$ sub-areas in Region A and $\hat{D}$ sub-areas in Region B.

Lemma 2.1. *(**Consistency with known sub-area numbers** C^* **and** D^*). Assume that* $\mathbb{E}(\mathbf{W}\mathbf{W}^{\mathbf{T}})$ *is of rank* C^* *with smallest absolute nonzero eigenvalue of at least* Λ_A*, and* $\mathbb{E}(\mathbf{W}^{\mathbf{T}}\mathbf{W})$ *is of rank* D^* *with smallest absolute nonzero eigenvalue of at least* Λ_B *. Assume further that* $\max(\mu_0, \mu_1, \sigma_0^2, \sigma_1^2) \leq d$ *for some* $d \leq \max(\log n/n, \log m/m))$*. Then, if there exists* $(2 + \varepsilon_A)\frac{ndCD}{\Lambda_A^2} < \tau_A$ *and* $(2 + \varepsilon_B)\frac{mdCD}{\Lambda_B^2} < \tau_B$ *for some* $\tau_A, \tau_B, \varepsilon_A, \varepsilon_B > 0$*, the output* $\hat{\mathcal{M}}_{(\hat{C},\hat{D})}$ *that maximizes function* (2.1) *is consistent to the true membership* $\mathcal{M}^*_{(C^*,\, D^*)}$ *underlying the latent community structure up to a permutation.*

Equivalently, let $\hat{S}_c$*,* $\hat{S}_d$ *be the estimated node sets for the subgraphs* G_c, G_d *(induced by* U_c *and* V_d*), respectively. Then* $\hat{S}_c \cap U_c$ *represents the nodes in* G_c *whose assignments can be guaranteed.* $\hat{S}_d \cap V_d$ *follows the same definition. With probability at least* $1 - \max(n, m)^{-1}$*, up*

to a permutation, we have

$$\sum_{c=1}^{C}\sum_{d=1}^{D}\left[1-\frac{\left|\,(\hat{S}_c\cap U_c)\bigotimes(\hat{S}_d\cap V_d)\,\right|}{|U_c|\,|V_d|}\right]\leq\tau_A^{-1}(2+\varepsilon_A)\frac{ndCD}{\Lambda_A^2}+\tau_B^{-1}(2+\varepsilon_B)\frac{mdCD}{\Lambda_B^2},$$

where $\bigotimes$ *denotes the edge set that connects two node sets on its left and right side.*

Theorem 2.2. ***(Consistency for grid-searched C, D)****. Let the sizes of subgraph pairs* $|U_c|\times|V_d|(\forall c=[C^*],d=[D^*])$ *be generated from a multinomial distribution with probabilities* $\pi=(\pi_1,\ldots,\pi_{C^*\times D^*})$*. Assume* $\exists\delta>0$*, such that*

$$\mu_1>\mu_0\frac{1+\delta}{1-\delta}\left(1+\sqrt{1+\frac{\pi_{\min}^2}{\pi_1^2+\cdots+\pi_{C^*\times D^*}^2}}\right),$$

then under conditions in Lemma 2.1 and tuning parameter $\lambda=0.5$*, the number of mis-assigned edges* N_{edge} *satisfies*

$$N_{\text{edge}}=o_p(n_{\min}*m_{\min})\text{as }|U|,|V|\to\infty,$$

where $n_{\min},m_{\min}$ *are the sizes of the smallest possible subgraphs in* U *and* V*, respectively.*

Theorem 2.3. ***(Convergence of Algorithm 1)****. Let* $\tilde{U}=\bigoplus_{c=1}^{\tilde{C}}U_c,\tilde{V}=\bigoplus_{d=1}^{\tilde{D}}V_d$ *be the partitions yielded by ratio-cut spectral clustering on* $\mathbf{W}_A$ *and* $\mathbf{W}_B$ *that maximizes function (1) with cluster numbers* $\tilde{C},\tilde{D}$*. Then* $\tilde{U},\tilde{V}$ *converge almost surely to the true community structure where false-positive edge assignments to each sub-bicluster are negligible.*

Proof. Proofs of Lemma 2.1 and Theorems 2.2 and 2.3 are provided in Appendix 2C.

In summary, the above results provide theoretical evidence that the solution of the proposed

objective function (2.1) and Algorithm 1 converge to the target community structure ($\{U_c, V_d\}$). Moreover, extensive simulation analyses in multiple settings with a wide range of different sample sizes demonstrate that SCCN can accurately reveal the true community structure with low false-positive and false-negative rates.

2.2.2.3 Determining $g(r)$ and λ

Determining $g(r)$**.** Following Efron (2012), we can choose $g(\cdot)$ to be a discrete distribution on thresholds $\{r_1, \ldots, r_p\}$. A simple example would be as follows. Suppose that the voxel-pair-level FDRs yielded by pre-selected thresholds r_1, r_2, and r_3 are 0.20, 0.10, and 0.05, respectively. We can then assign a higher probability mass to r_p that yields a lower FDR, for example, $g(r_1) = 0.1$, $g(r_2) = 0.3$, and $g(r_3) = 0.6$. In addition, r_1, r_2, and r_3 can be chosen from commonly used thresholds in MRI studies, such as $-\log(0.005)$ and $-\log(0.001)$ (W_{ij} is $-\log p_{ij}$ after screening, and r is a threshold for W_{ij}).

Selecting λ**.** As aforementioned, the tuning parameter λ adjusts the balance between the subgraph size and the connection density; it thus plays a critical role in our method. A large λ encourages large $|U_c|$ and $|V_d|$, whereas a small λ is stricter on the connection densities of (U_c, V_d) pairs. Essentially, the selection of λ is related to the network structure of β_{ij}. In practice, we have observed from many datasets that the coefficients $\beta_{ij} \neq 0$ usually exhibit a block model. To reflect this, we assume the following hierarchical model. Suppose there exists a non-random, latent hyperparameter $\boldsymbol{\beta} \in \mathbb{R}^{n \times m}$ with all nonzero elements. We can generate a bipartite similarity matrix $\eta \in \{0, 1\}^{n \times m}$ from a bipartite stochastic block model with blocks $\{(U_c, V_d)\}$ and the corresponding connection probabilities $\{\pi_{ij}\}$, such that $\eta_{ij} \sim \text{Bernoulli}(\pi_{ij})$ are independent of

each other, where

$$\pi_{ij} = \begin{cases} \pi_{cd}(\lambda) & i \in U_c, j \in V_d, \\ \pi_0(\lambda) & \text{otherwise.} \end{cases}$$

We select the λ value that maximizes the likelihood for this block model. In practice, the η_{ij} values are not directly observable, and we replace them by $\eta_{ij}(r_0) := I(w_{ij} > r_0)$. The log-likelihood function for λ is:

$$l_\lambda(\pi_{cd}, \forall c \in [C], d \in [D] | \eta_{ij}(r_0)) = \sum_{c,d} \sum_{(i,j) \in U_c \times V_d} \eta_{ij}(r_0) \log \pi_{cd} \quad + (1 - \eta_{ij}(r_0)) \log (1 - \pi_{cd}) \,.$$

To eliminate the arbitrariness in choosing the threshold r_0, we integrate the likelihood function with respect to r_0 over a prior distribution $g_0(r_0)$ determined by the method above. This yields the following criterion:

$$\lambda_{\text{optimal}} = \text{argmax}_\lambda \Big\{ \int \max_{U_c, V_d, \pi_{cd}} \quad l_\lambda^r \left(\pi_{cd}(\lambda), \forall c = [C], d = [D] \mid \eta_{ij}(r_0)\right) \; g_0(r_0) dr_0 \Big\}.$$

We formally present the procedure to select the tuning parameter λ in Algorithm 2. The overall complexity of the algorithm is $O(Knm)$, where K is a sufficient searching range for λ, $n = |U|$, and $m = |V|$.

Since the inference results between clinical groups across S subjects are captured in $\mathbf{W}$, the algorithm complexity no longer involves sample size S, indicating the scalability of SCCN for large datasets. In addition, clustering algorithms typically involve computing the first K eigenvectors of a potentially high-throughput similarity matrix. Our input similarity matrix $\mathbf{W}$ is sparse after applying screening and the spatial-contiguity constraints (usually only 0.2%–5.0% of edges are

non-zero entries after processing), which notably reduces computational expense. It is, however, worth noting that since our algorithm is based on a single region pair, the computational burden may become heavy when investigating multiple different pairs, especially when a whole-brain analysis is needed.

Algorithm 2 Grid search for λ.

procedure ALGORITHM
  for $0 \leq \lambda \leq 1$ **do**
    return $U = \bigoplus_{c=1}^{C} U_c$ and $V = \bigoplus_{d=1}^{D} V_d$ from *SCCN.partition* function in Algorithm 1
    for $r_0 = (r_0)_1$ to $(r_0)_q$ **do**
      Compute the log-likelihood: $l_\lambda(\hat{\pi}^{\text{MLE}}_{c\times d},\ \forall c = [C], d = [D] \mid \eta_{ij}(r_0))$
    end for
    Integrate the log-likelihood w.r.t. r_0:
    $l_\lambda = \sum_{i=1}^{p} L_\lambda \left(\hat{\pi}^{\text{MLE}}_{c\times d}, \forall c = [C], d = [D] |\ \eta_{ij}(r)\right)\ g((r_0)_i)$
  end for
  return $\hat{\lambda}$ that yields maximized l_λ
end procedure

2.2.3 Statistical inference of $\{(U_c, V_d)\}$ pairs

Recall that our ultimate goal is to extract a few most-densely connected subgraph pairs from $\{(U_c, V_d)\}$ based on the block partition $\{U_c, V_d : c \in [C], d \in [D]\}$ that we have already obtained at this point. A natural idea is to inspect each (U_c, V_d) pair and perform a statistical test on them with the alternative hypothesis that the subgraph $U_c \otimes V_d$ is unusually dense. Here, we derive a cluster-wise permutation test (Nichols and Holmes, 2002) with FWER control. The hypotheses are:

$\mathbb{H}_0$: G is a random graph (i.e., no dense subgraph $U_c \otimes V_d$ exists,

$\mathbb{H}_a$: At least one dense subgraph $U_c \otimes V_d$ exists.

More specifically, under $\mathbb{H}_0$, the connection density of $U_c \otimes V_d$ should be close to the density of sub-area pairs obtained by randomly shuffling edges in the bipartite graph. Built upon the minimum description length (MDL) principle proposed by Grünwald (2007), we devise the following MDL-based test statistic:

$$\mathrm{MDL}\left(U_c, V_d\right) = \log_2\left[\binom{n}{|U_c|}\binom{m}{|V_d|}\right] + \left(\frac{1-\mu_1^2}{2\ln 2} - L_\zeta\right)|U_c| \times |V_d|,$$

where μ_1 is the mean value of edge-wise test statistics ζ_{ij} for edges within $U_c \times V_d$, and $L_\zeta = -\int \phi(\zeta_{ij}) \log_2\left(\phi(\zeta_{ij})\right) d\zeta_{ij} + C$ is an information entropy measure based on the standard normal distribution ϕ for ζ_{ij}. Detailed derivations for the MDL-based test statistic and its connections to our inference goal are provided in Appendix 2C.4. We formally present the cluster-wise permutation test for each observed sub-area pair (U_c, V_d) in Algorithm 3. The number of permutations H in this algorithm can be determined based on the sample size, the targeted computational expense, and the precision of the test. For example, $H = 1000$. Compared to conventional multiple testing correction methods (e.g., FDR and FWER), the MDL-based cluster-wise permutation test returns suppressed false-positive findings and shows improved statistical power in real-data examples and simulations.

2.3 Simulations

In the simulation study, we probed whether SCCN can extract densely altered sub-area pairs with better performance compared to common existing methods. Specifically, we evaluated the performance from two perspectives. (i) Multivariate edge-level inference: whether extracted voxel pairs have a high true-positive rate (TPR) and low false-positive rate (FPR); (ii) network-level

Algorithm 3 MDL-based cluster-wise permutation test for each (U_c, V_d) pair

procedure ALGORITHM
  Compute $T^0_{c,d} = \mathrm{MDL}\,(U_c, V_d)$ for each (U_c, V_d) pair yielded with true covariate labels
  for $h = 1, \ldots, H$ **do**
    Permute covariate labels and obtain the new inference connectivity matrix $\mathbf{W^h}$
    Obtain $U^h = \bigoplus_{c=1}^{C} U_c^h$ and $V^h = \bigoplus_{d=1}^{D} V_d^h$ by substituting $\mathbf{W^h}$ in Algorithms 1 and 2
    return $T^h = \max\left(\mathrm{MDL}\left(\mathrm{U_c^h}, \mathrm{V_d^h}\right)\right)$
  end for
  Compute p-value for each observed (U_c, V_d) pair: $P_{c,d} = \frac{\sum I(T^h > T^0_{c,d})}{H}$
  return the significance of each observed (U_c, V_d) pair based on $P_{c,d}$ at a predetermined α-level
end procedure

inference: whether the extracted sub-areas contain maximal true-positive voxels, compared to other unextracted sub-areas.

2.3.1 Primary analysis

We first generated a bipartite graph $G = \{U, V\}$ to represent the brain connectome between two brain regions A and B for S subjects, where U corresponds to the voxel set in Region A, and V corresponds to that in Region B. We assume all S subjects share a common set of nodes after spatial normalization and registration, i.e., $(U^s, V^s) \equiv (U, V), \forall s \in [S]$. Next, we simulated covariates of interest $\{\mathbf{X}^1, \ldots, \mathbf{X}^S\}$ that contain clinical information of all S subjects. Lastly, we simulated the Fisher's z-transformation connectivity matrices $\{\mathbf{Z}^1, \ldots, \mathbf{Z}^S\}$ between regions A and B, where $\mathbf{Z}^s \in \mathbb{R}^{n \times m}, n = |U|, m = |V|$. Specifically, each element z_{ij}^s in $\mathbf{Z}^s$ was set to follow $\mathcal{N}(h(z_{ij}^s), \sigma^2)$, where $h(z_{ij}^s) = \mathbf{X}^s \beta_{ij}$ is location-specific within regions A and B.

In the following, we show the numerical settings under the above simulation framework:

1. For the two pre-defined brain regions of interest, we simulated $|U| = 900$ voxels in Region A and $|V| = 1600$ voxels in Region B. Within $|U|$ and $|V|$, we also randomly simulated three

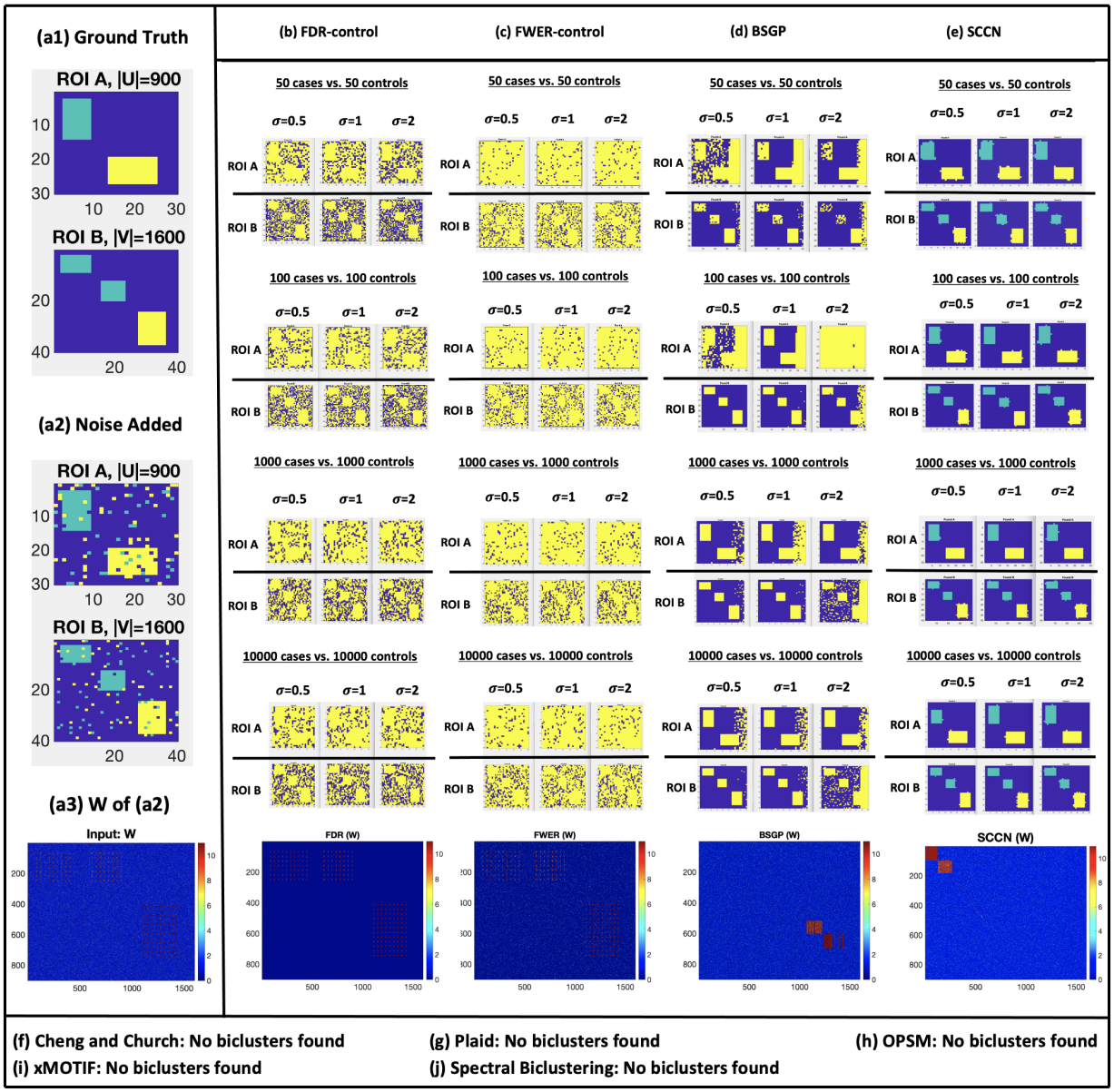


Figure 2.3: *A 2D visualization of performance by different methods. In (a1), the true spatial locations of sub-areas* U_1 *and* U_2 *are displayed in a simulated* 30×30 *grid space while* V_1*,* V_2*, and* V_3 *are displayed in a simulated* 40×40 *grid space. Sub-areas with the same color contain disease-related edges from A to B, i.e.,* (U_1, V_1)*,* (U_1, V_2)*, and* (U_2, V_3) *are the positive sub-area pairs. (a2) shows the scenario with false negative and false positive noises added to mimic the real vFC patterns in the brain connectome. (a3) shows the connectivity inference matrix* $\mathbf{W}$ *obtained based on (a2). (b)-(e) show the detected disease-related voxel pairs (again only regions with the same color form a pair) under different variances* σ *and sample sizes* S*. The last row shows the isomorphic graphs of (a3) with the extracted sub-area pairs pushed to the top when* $\sigma = 1$*. We highlight the voxels from the supra-threshold voxel pairs that were yielded by the FDR-control and FWER-control, and voxels in sub-area pairs that were extracted by BSGP and SCCN. Multiple testing with FDR-control and FWER-control tend to extract an excess of voxels with high false-positive error rates. BSGP better controls the error rates, but it extracts voxel pairs without differentiating the correct area-wise connections, i.e.,* (U_1, V_1)*,* (U_1, V_2)*, and* (U_2, V_3)*. In contrast, SCCN can simultaneously recover the spatially-contiguous sub-areas, respectively, in A and B, and reveal the correct disease-related vFC patterns. (f)-(j) show that no single differentially expressed sub-area pair was extracted by the biclustering algorithms listed.*

disease-related sub-area pairs (U_1, V_1), (U_1, V_2), and (U_2, V_3). The true spatial locations of these five sub-areas in the simulated 2d grid spaces are presented in Figure 2.3(a1). Not every possible pair $\{(U_c, V_d), c = [2], d = [3]\}$ was associated with the disease; only regions with the same color exhibited dysconnectivity from A to B. The sizes of these sub-area pairs were $|U_1||V_1| = 84 \times 70 = 5880$, $|U_1||V_2| = 84 \times 64 = 5376$, and $|U_2||V_3| = 96 \times 117 = 11\,232$. In addition, we included spatially isolated abnormal voxels as well as noise within regions A and B to mimic more realistic neural connectivity (Figure 2.3(a2)).

2. For the Fisher's z-transformation connectivity matrices $\{\mathbf{Z}^s, s \in S\}$, we set $h(z_{ij}^s) = \beta_0 + \beta_{ij,1}x_1^s + \beta_{ij,2}x_2^s + \beta_{ij,3}x_3^s$, where x_1^s and x_2^s store the age and sex information for subject s, and x_3^s represents their clinical status (x_3^s=1 if patient s has a mental disorder, and 0 for a healthy control.). In addition, while $\beta_{ij,1}$ and $\beta_{ij,2}$ are typically not spatially variant, $\beta_{ij,3}$ is considered brain-region specific:

$$\beta_{ij,3} = \begin{cases} 0.9, & \text{if} \quad (i,j) \in (U_1, V_1) \bigcup (U_1, V_2), \\ 0.13, & \text{if} \quad (i,j) \in (U_2, V_3), \\ 0, & \text{if} \quad (i,j) \in U/\{(U_1, V_1) \bigcup (U_1, V_2) \bigcup (U_2, V_3)\}. \end{cases}$$

3. To control standardized effect sizes, we set $\sigma^2 = 0.5, 1.0, 2.0$ in $\mathbf{Z}^s \sim \mathcal{N}(h(z_{ij}^s), \sigma^2)$. Additionally, four sample sizes, $S = 100$, 200, 2000, and $20,000$, were used, each with balanced healthy controls and patients. All settings with different (σ, S) were simulated for 1000 times to assess the variability of the TPR and FPR.

We implemented Algorithm 1 and 2 of SCCN to identify sub-area pairs from each simulated dataset, and we then applied Algorithm 3 to conduct cluster-wise inference on the sub-area pairs

detected. To assess the performance of the multivariate edge-wise inference, we considered two conventional multiple-testing controls (FDR and FWER). Specifically, we used the voxel-wise permutation test (with 1000 permutations) to control the FWER and the Benjamini–Hochberg procedure (with $q = 0.05$ as a cut-off) to control the FDR (Benjamini and Hochberg, 1995). To assess the accuracy of the cluster-wise performance, our goal was to compare true disease-related subgraphs $\{(U_c, V_d)\}$ with the estimated subgraphs $\{(\hat{U}_c, \hat{V}_d)\}$ produced by five commonly used biclustering algorithms (i.e., Cheng and Church, Plaid, OPSM, xMOTIF, and Spectral Biclustering (J. K. Gupta, 2013)).

The edge-wise inference results are presented in Table 1, and graph illustrations of the results is shown in Figure 2.3. For the edge-wise inference performance with all different (σ, S), SCCN outperforms the two traditional multiple testing correction methods (i.e., FDR and FWER control) in terms of TPR, while its ability to control the FPR falls in between the two. SCCN's relatively inferior performance in controlling the FPR (compared to sensitivity) can sometimes be impacted by the following disadvantage: in traditional multiple testing methods with universal thresholds, one false-positive finding corresponds to exactly one false-positive edge. However, SCCN detects altered edges by partitioning voxels within each ROI; therefore, one false-positive finding by SCCN corresponds to one false-positive voxel, say $v_i \in U_c$, which will lead to n false positive edges when V_d ($|V_d| = n$) is found to connect to $|U_c|$. The greater the size of V_d, the more false-positive edges will be yielded. Nonetheless, even with such a heavy penalty for detecting one false-positive voxel, SCCN still controls the FPR and shows better performance when jointly considering the TPR and FPR. More importantly, false-positive edges discovered by the traditional FDR and FWER correction approaches almost cover all within-ROI voxels, which leads to a substantial loss of spatial specificity when identifying covariate-related vFC patterns.

Sample Size: 50 cases vs. 50 controls				
Methods	σ	**0.5**	**1**	**2**
FDR-control	TPR	0.918 (0.229)	0.832 (0.314)	0.637 (0. 328)
	FPR	0.060 (0.007)	0.065 (0.053)	0.073 (0.158)
	Network Detection	No	No	No
FWER-control	TPR	0.994 (0.169)	0.893 (0.177)	0.842 (0. 179)
	FPR	0.762 (0.142)	0.764 (0.152)	0.771 (0.193)
	Network Detection	No	No	No
SCCN	TPR	1(0)	1(0)	1(0)
	FPR	0.087 (0.011)	0.093 (0.012)	0.096 (0.014)
	Network Detection	Yes	Yes	Yes
BSGP	TPR	0.946 (0.106)	0.893 (0.122)	0.667 (0.278)
	FPR	0.858 (0.258)	0.773 (0.173)	0.845 (0.439)
	Network Detection	Yes	Yes	Yes
Cheng and Church **Plaid** **OPSM** **xMOTIF** **Spectral Biclustering**	TPR	0(0)	0(0)	0(0)
	FPR	0(0)	0(0)	0(0)
	Network Detection	No biclusters detected		

Sample Size: 100 cases vs. 100 controls				
Methods	σ	**0.5**	**1**	**2**
FDR-control	TPR	0.958 (0.215)	0.892 (0.249)	0.746 (0. 427)
	FPR	0.065 (0.069)	0.069 (0.048)	0.070 (0.164)
	Network Detection	No	No	No
FWER-control	TPR	0.994 (0.157)	0.989 (0.182)	0.842 (0. 163)
	FPR	0.767 (0.139)	0.762 (0.140)	0.761 (0.187)
	Network Detection	No	No	No
SCCN	TPR	1(0)	1(0)	1(0)
	FPR	0.081 (0.010)	0.082 (0.012)	0.097 (0.013)
	Network Detection	Yes	Yes	Yes
BSGP	TPR	1 (0.105)	1 (0.120)	0.993 (0.262)
	FPR	0.844 (0.228)	0.764 (0.172)	0.940 (0.369)
	Network Detection	Yes	Yes	Yes
Cheng and Church **Plaid** **OPSM** **xMOTIF** **Spectral Biclustering**	TPR	0(0)	0(0)	0(0)
	FPR	0(0)	0(0)	0(0)
	Network Detection	No biclusters detected		

Sample Size: 1000 cases vs. 1000 controls				
Methods	σ	**0.5**	**1**	**2**
FDR-control	TPR	0.957 (0.209)	0.878 (0.287)	0.789 (0. 363)
	FPR	0.061 (0.007)	0.061 (0.058)	0.071 (0.159)
	Network Detection	No	No	No
FWER-control	TPR	0.993 (0.169)	0.904 (0.180)	0.823 (0. 187)
	FPR	0.749 (0.140)	0.727 (0.145)	0.754 (0.176)
	Network Detection	No	No	No
SCCN	TPR	1(0)	1(0)	1(0)
	FPR	0.081 (0.008)	0.087 (0.010)	0.081 (0.012)
	Network Detection	Yes	Yes	Yes
BSGP	TPR	0.968 (0.106)	0.903 (0.122)	0.875 (0.278)
	FPR	0.374 (0.438)	0.405 (0.360)	0.737 (0.447)
	Network Detection	Yes	Yes	Yes
Cheng and Church **Plaid** **OPSM** **xMOTIF** **Spectral Biclustering**	TPR	0(0)	0(0)	0(0)
	FPR	0(0)	0(0)	0(0)
	Network Detection	No biclusters detected		

Sample Size: 10000 cases vs. 10000 controls				
Methods	σ	**0.5**	**1**	**2**
FDR-control	TPR	0.959 (0.229)	0.873 (0.357)	0.771 (0. 391)
	FPR	0.060 (0.007)	0.061 (0.061)	0.069 (0.160)
	Network Detection	No	No	No
FWER-control	TPR	0.992 (0.174)	0.907 (0.176)	0.819 (0. 183)
	FPR	0.753 (0.145)	0.734 (0.145)	0.759 (0.157)
	Network Detection	No	No	No
SCCN	TPR	1(0)	1(0)	1(0)
	FPR	0.081 (0.007)	0.091 (0.007)	0.087 (0.011)
	Network Detection	Yes	Yes	Yes
BSGP	TPR	0.972 (0.208)	0.884 (0.122)	0.873 (0.290)
	FPR	0.574 (0.426)	0.368 (0.499)	0.679 (0.407)
	Network Detection	Yes	Yes	Yes
Cheng and Church **Plaid** **OPSM** **xMOTIF** **Spectral Biclustering**	TPR	0(0)	0(0)	0(0)
	FPR	0(0)	0(0)	0(0)
	Network Detection	No biclusters detected		

Table 1: Simulation results. The four sub-tables show the inference results given different sample sizes and variances, where TPR and FPR correspond to the edge-wise true positive rate and false positive rate. Network detection results indicate whether the algorithm can successfully extract the correct connection patterns between disease-related sub-area pairs.

Regarding the network-level inference performance, all common biclustering methods failed to detect any positive biclusters (differentially expressed sub-area pairs) except for BSGP. However, BSGP nonetheless failed to ensure spatial contiguity, and the precise connection between the extracted sub-areas was not correctly revealed. That is, unlike the results yielded by SCCN (Figure 2.3(e)), BSGP (Figure 2.3(d)) could not effectively differentiate between yellow and blue clusters. In comparison, SCCN shows outstanding network-level performance for detecting community structures and incorporating spatial contiguity.

2.3.2 Negative control analysis

We further performed a negative control analysis to evaluate the FPR of our method. We consider a scenario in which the connections between a pre-selected ROI pair are unrelated to a clinical condition of interest. We generated $|U| = 900$ and $|V| = 1600$ voxels in regions A and B. We distinguished the patient and control groups as 1 and 0, but since there were no abnormal sub-area pairs $\{(U_c, V_d)\}$ across groups, we simply set the connectivity matrices $\mathbf{Z}^s \sim \mathcal{N}(0, \sigma^2)$ over the entire regions for all S subjects. Based on $\mathbf{Z}^s$, we obtained the inference matrix $\mathbf{W^0}$ across clinical groups. Since the network detection was validated to be scalable to different sample sizes and sample variances, we evaluated the configuration $(S = 1000, \sigma = 1)$ as a proof of concept. Finally, we implemented SCCN on $\mathbf{W^0}$. Since the false positive voxel pairs tended to be distributed randomly, no sub-area pairs were significant. Therefore, the sub-area-level false positive findings were 0. The edge-wise FPR (supra-threshold voxel-pairs) among 1000 iterations was $6.82 \times 10^{-5}(\text{std. } 1.29 \times 10^{-5})$, which is consistent with the pre-determined alpha level ($\mathbb{E}(p) = 0.00005$). We have provided a graph visualization of these results in Appendix 2F.

In summary, we have shown that the sub-area detection is not affected by different values of variance σ^2, sample size S, or other sources of noise. SCCN also yields vFC patterns with high sensitivity and low FPRs. The spatial-contiguity constraints allow positive edges to borrow strengths from each other within a data-driven sub-area; sensitivity is thus notably increased. Data-driven sub-areas with these constraints can also exclude false-positive edges that bridge voxels that are randomly scattered in ROIs. False-positive findings are therefore largely suppressed. In addition, the jointly improved sensitivity (and thus statistical power) and control of the FPR yield almost identical voxel sets across all simulated datasets. Replicability is hence remarkably improved.

2.4 Real data application

In this section, we apply SCCN to two real datasets to investigate the voxel-level altered connections under different clinical settings. Dataset 1 contains 3269 subjects from a nicotine-addiction study using fMRI data collected from the UK Biobank database. Dataset 2 includes 330 subjects from a schizophrenia (SZ) research study using fMRI data collected in Baltimore, MD.

2.4.1 Nicotine-addiction research study

2.4.1.1 Sample characteristics

Our primary dataset consists of 3269 individuals from the UK Biobank database, including 1353 current smokers (M/F: 737/616, age: 48.6 ± 15.3) and 1916 previous light smokers (M/F: 1187/729, age: 32.9 ± 18.1). Additional information on the selection of these 3269 subjects is provided in Appendix 2E.1. Specifically, we define *current smokers* as participants who currently

smoke more than ten cigarettes per day, indicating nicotine addiction [2]. Conversely, we define *previous light smokers* as individuals who had tried only a few cigarettes in the past but are not currently addicted to nicotine products, serving as controls [3]. For detailed information on fMRI imaging acquisition and preprocessing procedures, please refer to Appendix 2E.2.

2.4.1.2 Clinical background

Abundant literature shows that the basal ganglia (BG), hippocampus (Hippo), and insular gyrus (Ins) play important roles in nicotine addiction (Ersche et al., 2011; Gaznick et al., 2014; McClernon et al., 2016). We therefore intend to look into the disrupted connectivity patterns between these three bilateral ROIs, resulting in a total of 12 pairs. To maintain conciseness, we present the results for the (left BG, left Ins) pair in the main text, while the remaining 11 cases are provided in Appendix 2E.4. By investigating the altered vFC patterns across different clinical groups, we aim to gain insights into the underlying neurological mechanisms of nicotine dependence and ultimately assist smokers in resisting nicotine cravings.

We labeled the left BG and left Ins using the Brainnetome Atlas(Fan et al., 2016)(left BG: 2345 voxels; left Ins: 1762 voxels). For each subject, we calculated the vFC matrix between the left BG and left Ins, with each entry representing a Fisher's z-transformed Pearson correlation coefficient. Next, we calculated the population-level statistical inference matrix $\mathbf{W}_{2345\times1762}^{(\mathrm{BG_{left}},\mathrm{Ins_{left}})}$ across all subjects while adjusting for age, sex, site, educational level, and Body Mass Index (BMI). More details regarding the selection of nuisance covariates can be found in Appendix 2E.3. Applying SCCN and the MDL-based test, we extracted abnormal sub-area pairs from

[2] ACE touchscreen question "About how many cigarettes do you smoke on average each day? "
[3] ACE touchscreen question "In the past, how often have you smoked tobacco?"

$\mathbf{W}^{(\mathrm{BG}_{\mathrm{left}},\mathrm{Ins}_{\mathrm{left}})}$ with spatial-contiguity constraints. Lastly, we compared these results with those obtained from comparative methods.

2.4.1.3 Network-level results

Each entry in the inference matrix $\mathbf{W}^{(\mathrm{BG}_{\mathrm{left}},\mathrm{Ins}_{\mathrm{left}})}$ is endowed with a $-\log p$ value testing the vFC difference between current smokers and previous light smokers (Figure 2.5(1)). Implementing Algorithm 2 returned the MLE $\hat{\lambda} = 0.75$. Given the estimated $\hat{\lambda}$, Algorithm 1 returned the number of clusters $\hat{C} = 306, \hat{D} = 210$ for $\mathbf{W}^{(\mathrm{BG}_{\mathrm{left}},\mathrm{Ins}_{\mathrm{left}})}$. The MDL-based test returned six abnormal sub-area pairs, which are marked in red in Figure 2.5(2). A 3D demonstration of the detected sub-area pairs from $\mathbf{W}^{(\mathrm{BG}_{\mathrm{left}},\mathrm{Ins}_{\mathrm{left}})}$ is shown in Figure 2.5(a)–(e) (with a significance level of 0.05 selected for the MLD-based permutation test). All extracted sub-area pairs show well-organized topological structures. Results indicate that the majority of aberrant vFC patterns from $\mathbf{W}^{(\mathrm{BG}_{\mathrm{left}},\mathrm{Ins}_{\mathrm{left}})}$ are gathered between the medial inferior part of the left basal ganglia and the left insula.

2.4.1.4 Biological interpretation of detected sub-areas

The detected sub-areas consist of several locations that are believed to be frequently associated with nicotine addiction, including the medial inferior part of the basal ganglia and the posterior insula. We also observed decreased connectivity within these regions in current smokers, which aligns with the previous medical discovery that decreased resting-state functional connectivity is correlated with increased nicotine-addiction severity (Fedota and Stein, 2015; Sutherland and Stein, 2018). The incorporated spatial-contiguity constraints help unfold the sub-areas within the BG, Hippo, and Ins, which maximally cover addiction-related vFC. These

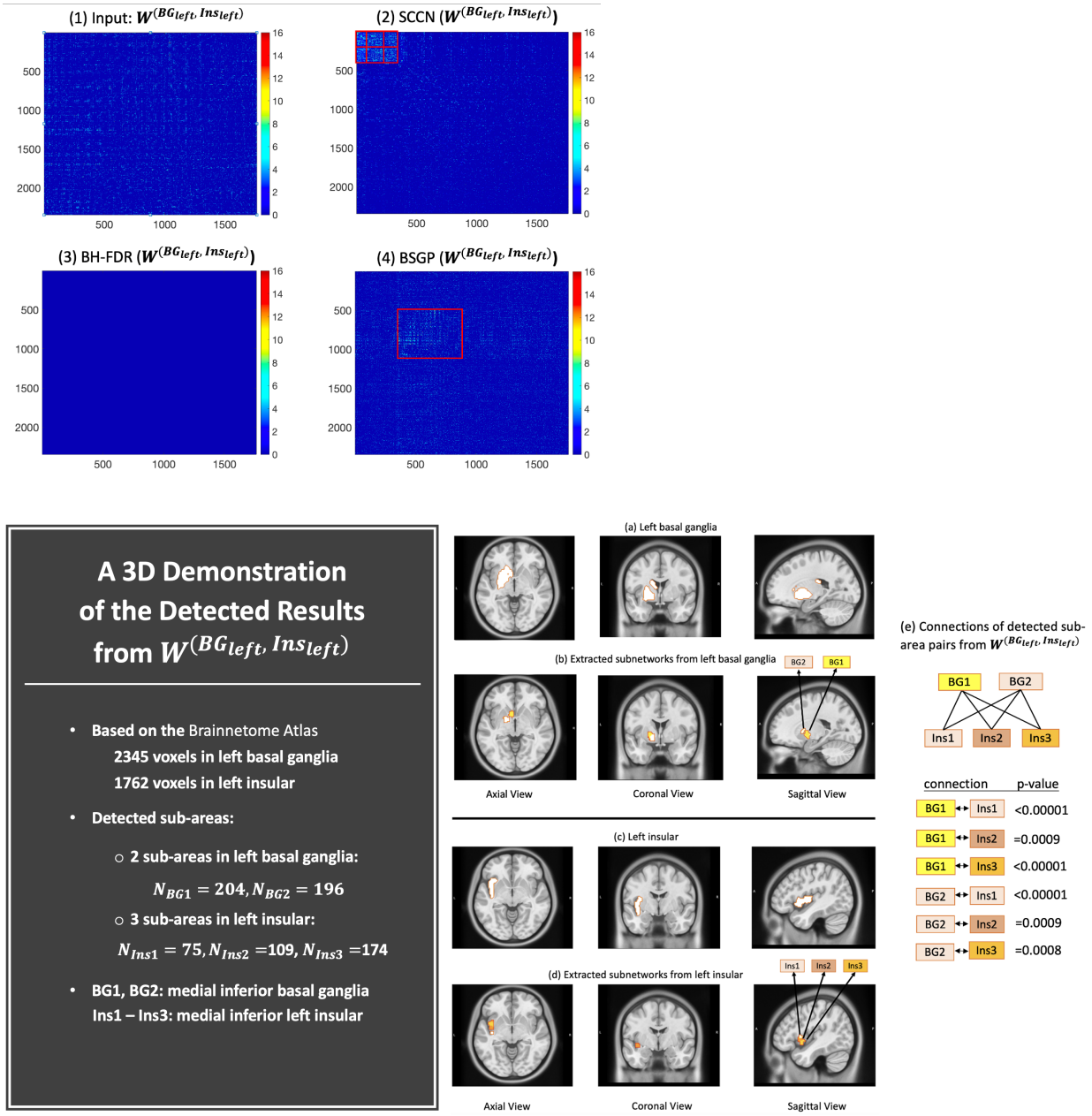


Figure 2.5: *Detected sub-area pairs from a nicotine-addition study. (1) A heatmap of* $\mathbf{W}^{(\mathrm{BG_{left}}, \mathrm{Ins_{left}})}$*: rows and columns correspond to voxels from the left basal ganglia and the left insula, respectively. (2) Results yielded by SCCN: altered sub-area pairs that pass the MDL-based permutation test are highlighted in red boxes. (3) Results yielded by BH-FDR: The hypothesis testing error measure was set to be* $q = 0.05$ *as a cut-off. No sub-area pairs were detected. (4) Results yielded by BSGP: only one positive yet much less dense sub-area pair was detected. The detected sub-area pair also lack spatial contiguity and specificity. (a)-(d) shows the 3D demonstration of the 6 detected altered sub-areas from* $\mathbf{W}^{(\mathrm{BG_{left}}, \mathrm{Ins_{left}})}$*. (a)-(e) show a 3D demonstration of the detected results from* $\mathbf{W}^{(\mathrm{BG_{left}}, \mathrm{Ins_{left}})}$*. Based on the p-values from the MDL-based permutation test shown in (e), most positive sub-area pairs are located in the medial inferior part of left basal ganglia and left insula.*

novel findings improve the spatial specificity of addiction-related locations in the three brain regions and may lead to future guidance for resisting the urge to use nicotine products.

2.4.1.5 Comparisons with existing methods

For comparison purposes, we again performed the BH-FDR correction edge-wisely and BSGP cluster-wisely on $\mathbf{W}^{(\mathrm{BG}_{\mathrm{left}},\mathrm{Ins}_{\mathrm{left}})}$. By first conducting an initial edge-wise significance test across the current and previously light smoker groups, only 7.29% of the edges were found to be significant ($p < 0.005$). However, no edges showed significance after applying BH-FDR correction with $q = 0.01$ (Figure 2.5(3)). When applying BSGP to $\mathbf{W}^{(\mathrm{BG}_{\mathrm{left}},\mathrm{Ins}_{\mathrm{left}})}$, only one abnormal sub-area pair was detected (Figure 2.5(4)), with 49.5% edges of $p > 0.005$ included in the detected pair, compared to 3.12% yielded by SCCN. In comparison to the two existing methods, SCCN yields much more densely altered vFC contained in spatially contiguous sub-area pairs with strong topological structures.

2.4.2 Schizophrenia research study

2.4.2.1 Sample characteristic

Our primary dataset contains 330 individuals, including 148 SZ patients (M/F 84/64, age 37.5 ± 14.4) and 182 healthy controls (M/F 80/102, age 37.0 ± 16.1). The participants were required for a large ongoing study of the effects of cognitive deficits in SZ. Specifically, the study probed how cognitive deficits contributed to functional disability in SZ patients and how they were related to altered functional networks that serve cognition. All subjects were assessed at local research centers in the greater Baltimore area between 2004 and 2016 using uniform recruitment criteria,

and neurological and clinical assessments. Detailed information about participant demographics, the recruitment process, imaging acquisition, and fMRI preprocessing procedures can be found in Appendix 2D.1.

2.4.2.2 Salience network disrupted connectivity

Clinical background The salience network, which is mainly composed of the bilateral insula and cingulate cortices, is related to several core SZ symptoms. A vast amount of literature in neuroimaging research suggests that the connectivity in the salience network is disturbed during information processing in SZ patients (Palaniyappan et al., 2012). We therefore intend to focus on the bilateral insula and cingulate cortices and study the schizophrenic-altered vFC patterns between them. Specifically, we want to extract schizophrenic-impacted edges that connect voxels from spatially coherent sub-areas within the insula to those within the cingulate cortex. This data-driven extraction of sub-areas caused by vFC abnormality in SZ may provide insights for more effective clinical treatments (e.g., by transcranial magnetic stimulation or deep-brain-stimulation therapies). We labeled the bilateral insula and cingulate cortices based on the Brainnetome Atlas (Fan et al., 2016) (left insula: 1762 voxels; right insula: 1577 voxels; cingulate cortex: 5768 voxels). We applied SCCN and the MDL-based test to the edge-wise connectivity inference matrices $\mathbf{W}^L_{1762\times5768}$ and $\mathbf{W}^R_{1577\times5768}$ for the (left Ins, cingulate) and (right Ins, cingulate) ROI pairs respectively while adjusting for age and sex. $\mathbf{W}^L$ and $\mathbf{W}^R$ were obtained by the same computational procedures as in Dataset 1. Lastly, we compared the detection results with those obtained by comparative methods.

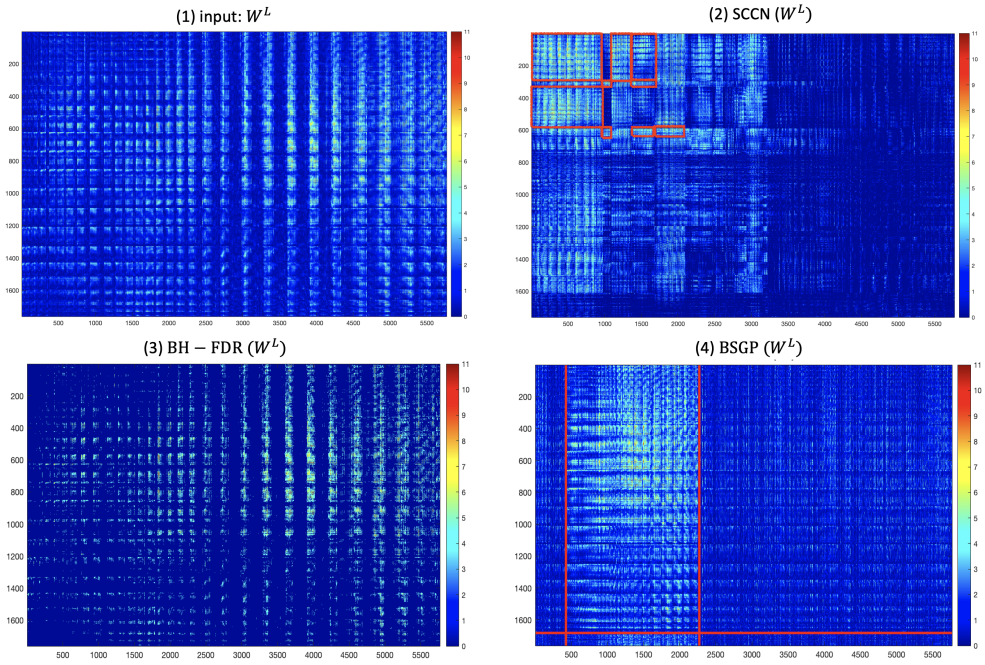


Figure 2.6: *Detected sub-area pairs in salience network from a schizophrenia study (2D). (1) A heatmap of* $\mathbf{W}^L$*: rows and columns correspond to the voxels from the left insula and the cingulate cortex, respectively. A hotter entry indicates a more differentially expressed voxel pair between clinical groups adjusted for other covariates. (2) Results yielded by SCCN: positive sub-area pairs that pass the MDL-based permutation test are highlighted in red boxes. There are many edges with small p-values outside the red boxes (e.g., in the bottom left corner) because they are not spatially contiguous to those inside the boxes, and are automatically excluded by SCCN. (3) Results yielded by BH-FDR: with* $q = 0.05$*, no sub-area pairs were detected. (4) Results yielded by BSGP: only one informative yet much less dense sub-area pair was detected. The detected sub-area pair was also lack of spatial contiguity and specificity.*

Network-level results Each element in the vFC inference matrix $\mathbf{W}^L$ is $\mathbf{W}^L_{ij} = -\log(p^L_{ij})$, where p^L_{ij} is the p-value testing the case-control vFC difference for the (i, j) pair between the left insula and cingulate cortex (Figure 2.6(L1)). We then perform screening on $\mathbf{W}^L$ using a pre-selected threshold (e.g., $p = 0.05$): $\mathbf{W}^L_{ij} = (\mathbf{W}^L)_{ij} \cdot I\left((\mathbf{W}^L)_{ij} \leq -\log(0.05)\right)$. The post-screened inference matrix $\mathbf{W}^L$ can effectively exclude most non-informative false-positive edges while maintaining a high proportion of true-positive edges (Fan and Lv, 2008; Li et al., 2012a). Similar settings apply to $\mathbf{W}^R$ (Figure 2.6(R1)). Implementing Algorithm 2 returned a maximum-likelihood estimation (MLE) of $\hat{\lambda}_L = 0.625$ for $\mathbf{W}^L$ and $\hat{\lambda}_R = 0.75$ for $\mathbf{W}^R$. Given the estimated $\hat{\lambda}$, Algorithm 1 returned the number of clusters $\hat{C}_L = 135, \hat{D}_L = 107$ for $\mathbf{W}^L$, and

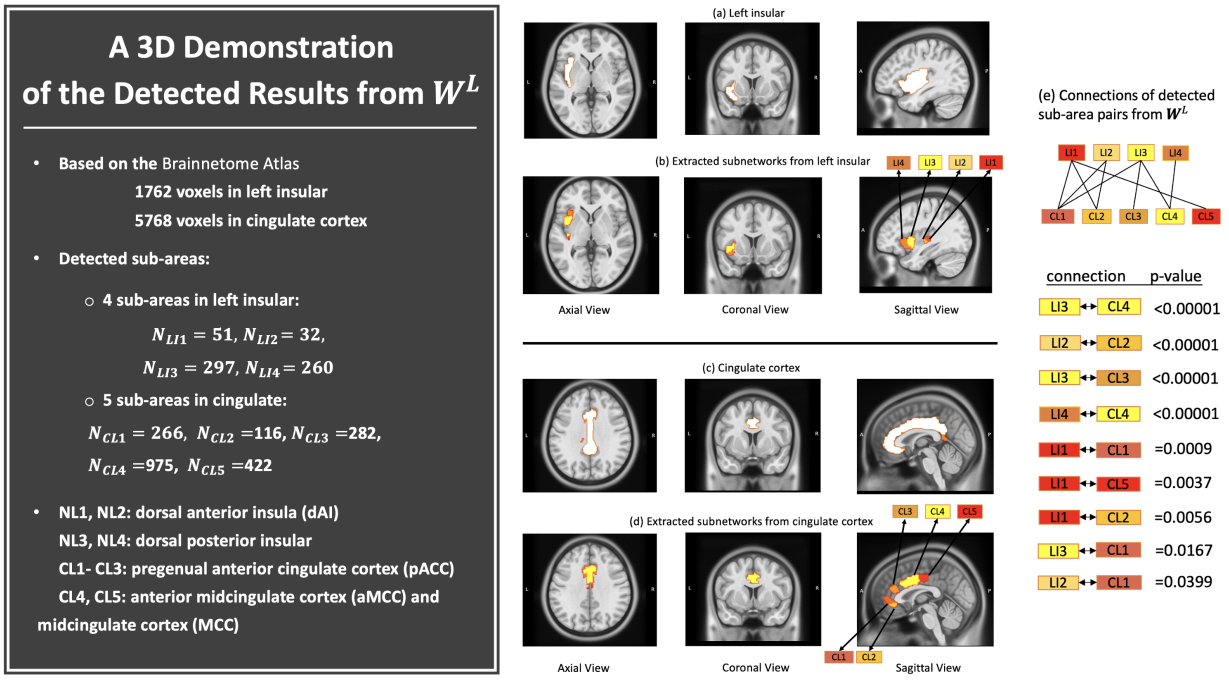


Figure 2.7: *Detected sub-area pairs in salience network from a schizophrenia study (3D). Let LI_i be the i-th disrupted sub-area detected from the left insula that is connected to the j-th sub-area from the cingulate cortex, CL_j. Let N_{LI_i} denote the number of voxels in sub-area LI_i, and similarly N_{CL_j} for CL_j. (a)(c) show the images of the original left insular and cingulate cortex; (b) shows the SZ-affected sub-areas in the left insula that are connected to those in the cingulate cortex highlighted in (d); (e) shows the architecture of interconnections between the detected sub-areas from $\mathbf{W}^L$ and the associated p-values from the MDL-based permutation test. A 3D demonstration of the detected results from $\mathbf{W}^R$ is provided in Appendix 2D.2.*

$\hat{C}_R = 225, \hat{D}_R = 226$ for $\mathbf{W}^R$. The MDL-based test returned nine abnormal sub-area pairs for $\mathbf{W}^L$ and ten abnormal sub-area pairs for $\mathbf{W}^R$ (marked in red in Figure 2.6(L2) and (R2)). A 3D demonstration of the detected results from $\mathbf{W}^L$ is shown in Figure 2.7 (using a significance level of 0.05 from the MDL-based permutation test). Information regarding the precise sizes, p-values, and locations is also specified in Figure 2.7. All extracted sub-area pairs show well-organized topological structures. Overall, the aberrant vFC patterns from $\mathbf{W}^L$ are gathered between the dorsal insula and anterior cingulate cortex (ACC). Detailed detection results for $\mathbf{W}^R$ are provided in Appendix 2D.2.

Biological interpretation of detected sub-areas The detected sub-areas consist of several well-known brain regions that are believed to be frequently associated with SZ disorder, including, most remarkably, the anterior insula (AI) and ACC. Emotions that most strongly engage the AI, such as anger and fear, are those that SZ patients tend to have the most difficulty recognizing (Wylie and Tregellas, 2010). Furthermore, the densities of neurons, axons, and synapses are found to be abnormal in the ACCs of people with SZ (Arnold and Trojanowski, 1996). All of the aberrant edges detected showed decreased or equivalent connections in SZ patients. This aligns with medical findings that SZ is a "dysconnectivity" disorder with primarily reduced FC across the salience network (Lynall et al., 2010), although medication effects cannot be completely ruled out. The imposed spatial-contiguity constraints help unfold brain sub-areas of the bilateral insula and cingulate cortices that maximally cover disease-related vFC. These novel findings improve the spatial specificity of SZ-related dysconnectivity in the well-known salience network and may lead to guidance for future treatments.

Comparisons with existing methods For comparison purposes, we performed the Benjamini–Hochberg FDR (BH-FDR) correction edge-wisely and a commonly used biclustering algorithm, bipartite spectral graph partitioning (BSGP), cluster-wisely. By first conducting an initial correlation analysis between vFC and schizophrenic status, 17.84% of the edges in $\mathbf{W}^L$ were found to have $p < 0.005$ significance, where $p = 0.005$ is a commonly used yet uncorrected threshold in neuroimaging studies (Derado et al., 2010). After applying BH-FDR correction, 9.45% of the edges were found to be significant using the threshold of $q = 0.01$ (Figure 2.6(L3)), and no community structure was revealed. For $\mathbf{W}^R$, 13.50% of edges had p-values less than 0.005, and only 3.61% significant edges were found after BH-FDR correction with $q = 0.01$ (Figure 2.6(R3));

again, no community structure was found in $\mathbf{W}^R$. When applying BSGP to both $\mathbf{W}^L$ and $\mathbf{W}^R$, only one abnormal sub-area pair was detected (Figure 2.6(L4) and (R4)), with more than 36.80% edges of $p > 0.005$ included compared to SCCN. In comparison to the existing methods, SCCN yields much more densely schizophrenia-associated vFC contained in spatially contiguous sub-area pairs with stronger topological structures.

2.4.2.3 Temporal-thalamic disrupted connectivity

In contrast to the reduced salience network connections in SZ patients, many studies have shown that SZ patients have greater thalamic connectivity with multiple sensory-motor regions, including, most remarkably, the temporal gyrus (Cetin et al., 2014; Ferri et al., 2018). More specifically, thalamus to middle temporal gyrus connectivity was positively correlated with many core SZ features, such as hallucinations and delusions. We therefore aim to use SCCN to identify some novel findings between the middle temporal gyrus on the right hemisphere and the bilateral thalamus in SZ patients. Based on the Brainnetome Atlas, there are 3566 voxels in the right middle temporal gyrus (labeled 82, 84, 86, and 88) and 3275 voxels in the bilateral thalamus (labeled 231–246). We computed the vFC connectivity inference matrices $\mathbf{W}_{3566\times 1727}^{(\mathrm{Tem}_{\mathrm{right}},\mathrm{Tha}_{\mathrm{left}})}$ and $\mathbf{W}_{3566\times 1548}^{(\mathrm{Tem}_{\mathrm{right}},\mathrm{Tha}_{\mathrm{right}})}$ between clinical groups and then implemented SCCN. Due to limited space here, we provide the results for the selections of all parameters and densely altered sub-area pairs in Appendix 2D.3.

2.5 Discussion

Psychiatric and neurological disorders are often associated with a disrupted brain connectome. To improve the spatial specificity and sensitivity for detecting a disease-impacted brain connectome, in this work, we focused on voxel-level connectivity network analysis. We developed statistical models focusing on extracting abnormal voxel pairs from a region pair of interest, which can be further extended to whole-brain connectome analysis. We have attempted to simultaneously address the challenges of a controlled FPR for multiple voxel-pair testing and the spatial-contiguity constraints for vFC analysis.

In addition, the brain parcellation to extract sub-areas is usually based on commonly used brain atlases (e.g., Brodmann's map or the International Consortium for Brain Mapping), and these were built on comprehensively studied cortical anatomy, such as complex gyro-sulcal folding patterns. Different regions blocked by gyri and sulci tend to show differential neurobiological structures and functions, and these atlases can thus serve as a good foundation to investigate sub-area community structures. However, to further overcome the limitation of using existing brain parcellations, one can consider combining any extracted spatially adjacent sub-areas from a pair of spatially adjacent regions if the combination is statistically coherent and biologically meaningful.

The centerpiece of our proposed method is the identification of sub-area pairs containing an unusually high density of phenotype-related voxel pairs. By leveraging this high density, we can effectively control the FPR by excluding isolated false-positive edges, and we thus greatly reduce the number of false-positive nodes. We have therefore improved the spatial specificity of extracted disease-related patterns at a voxel level. Herein, we have proposed a new non-parametric objective

function to achieve this goal, and this has been implemented with efficient algorithms. We also developed inference methods to assess the statistical significance of each sub-area pair extracted.

The biological findings from our data example are novel; SCCN revealed vFC connectome patterns for schizophrenia within the well-known salience network. We discovered that the malfunction of salience network connectivity is mainly driven by disrupted connections between the dorsal insula and anterior cingulate cortex instead of the omnibus region-level findings. We further validated our findings through extensive simulations and showed that our methods could improve sensitivity with a controlled FPR while retaining spatial contiguity.

In summary, SCCN provides a new toolkit for vFC analysis with improved spatial resolution and specificity while preserving a well-controlled false-positive error rate. Therefore, the findings from SCCN can be translated into more effective potential treatments for brain disorders. Since the input data of SCCN is voxel-pair-level inference results, it is applicable to all connectivity measures and data modalities where valid statistical inference can be performed (e.g., white-matter tractography). SCCN may also provide a promising strategy for whole-brain connectome voxel-pair network analysis. All sample code can be found at https://github.com/TongLu-bit/DecodingNetwork-SCCN.

Chapter 3: High Dimensional Multiple Imputation (HIMA))

3.1 Introduction

Neuroimaging data is crucial for studying the structure and function of the brain, providing valuable insights into various neurological disorders and cognitive processes. In brain imaging research, missing data however frequently occurs due to limited image acquisition and susceptibility artifacts, which cause signal loss and spatial distortion in the images (Mulugeta et al., 2017). Figure 3.1 shows an example of the spatial distribution of missing voxels in a magnetic resonance imaging (MRI) dataset. Despite advancements in statistical techniques for processing imaging data, the issue of neuroimaging missingness remains inadequately addressed. The missingness hinders accurate analysis and interpretation of findings. For instance, missing data can introduce bias, reduce statistical power, and limit the generalizability of results. We are thus motivated to propose an applicable and robust multivariate imputation technique specifically designed for high-dimensional neuroimaging data.

The commonly used strategies to handle incomplete data include (i) complete data analysis (ii) single imputation (iii) multiple imputation. In neuroimaging studies, the missing data problem is commonly addressed by (i), i.e., omitting missing voxels from the analysis. Simply omitting these voxels may risk excluding brain regions of particular study interest and may be costly in terms of other spatial coverage, especially along cortical boundaries. Consequently, the risk for

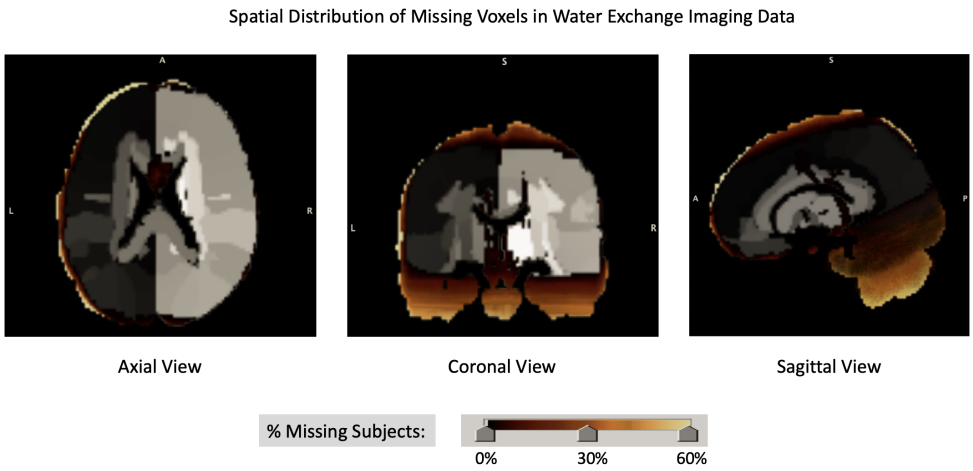


Figure 3.1: *Neurovascular water exchange imaging data was collected from a schizophrenia research ($N = 58$). Shown from different cross-sectional views, grey voxels contain non-missing data while the gold color scale indicates the percentage of subjects with missing data. Darker-colored voxels contain missing data from a smaller proportion of subjects and vice versa.*

Type II errors may also increase. On the other hand, simple imputation improves upon (i) and substitutes a single value for each missing value (e.g., mean/mode substitution). While quick fixes like simple imputation may be effective in some cases, they usually introduce bias into the data (e.g., artificial reductions in variability) and lead to overly precise results without accounting for any uncertainty. Schafer (1997) and Rubin (2004) addressed this by developing multiple imputation (MI) techniques that can incorporate uncertainty about the unknown missing values. MI replaces each missing value with a set of plausible values that are imputed based on two factors: (a) the observed values for a given subject; (b) the relations observed in the data for other subjects.

Statistical literature has established that MI techniques are sound. However, the application of MI is limited primarily because of the computational cost and tractability. For example, **M**ultivariate **I**mputation by **C**hained **E**quations (MICE) is one of the most commonly used MI toolboxes (Enders, 2022; Royston and White, 2011; Vaden Jr et al., 2012). Given approximate normal data, MICE specifies the inverse-Wishart (IW) distribution as a conjugate prior distribution

for the normal covariance $\boldsymbol{\Sigma}$. Sampling a large $\boldsymbol{\Sigma}_{p\times p}$ matrix (e.g., $p = 1000$) can become computationally unstable or intractable, which may lead to inaccurate subsequent draws of data and eventually inaccurate imputation results. In addition, sampling $\boldsymbol{\Sigma}_{p\times p}$ involves matrix inversion, which requires a computation cost of $\mathcal{O}(p^3)$. The cubic time complexity will be further compounded by the number of sampling iterations and the number of total imputed datasets needed. In Figure 3.2, we plotted the time taken to impute missingness in a real MRI dataset from a schizophrenia study using the MICE package in R. It is noticeable that the computational time grows exponentially as the number of voxels rises. For a typical brain region with hundreds of variables, it can take MICE thousands of hours to impute, which is not quite computationally feasible.

As indicated by Figure 3.2, the complexity and tractability can be significant computational bottlenecks for high-dimensional imaging data. To address this challenge, we propose a new **H**igh-**d**imensional **M**ultiple Imput**a**tion (HIMA) method, which is particularly applicable to impute missingness in high dimensional data. HIMA adopts the commonly used Bayesian framework

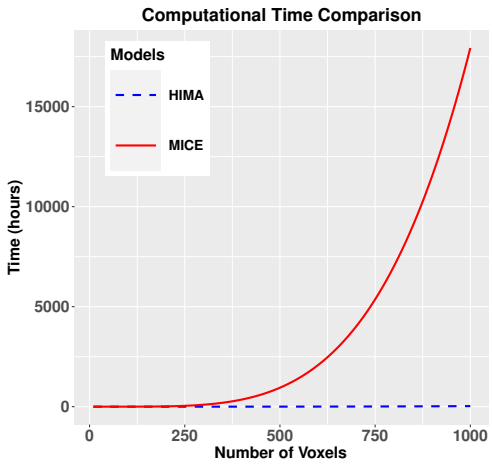


Figure 3.2: *Running time (in hours) against the number of voxels using MICE and HIMA (proposed model). The number of imputed datasets is set to be 10; within each dataset, there are 15 iterations. Both MICE and HIMA were implemented using a single computational node with 128 GB memory and 8 Core CPU.*

under the framework of Markov chain Monte Carlo (MCMC) for missing data imputation, e.g., MICE. Specifically, MICE implements the same imputation step for missing entries by considering a joint multivariate normal (MVN) model while modifying the posterior step by sampling the covariance matrix from the posterior mode. The mode represents the most probable draws from the posterior distribution, which is also essentially identical to the ML estimates of the likelihood functions (Schafer (1997)). Exploiting the posterior mean relieves ultra-high computational burdens from $\mathcal{O}(p^3)$ to $\mathcal{O}(p)$ per iteration, and avoiding large covariance matrix sampling also helps obtain more accurate imputation results. We have adopted a posterior mode estimator with good asymptotic properties that perform well for $n \ll p$, where n is the sample size.

The contributions of this article are three-fold. Firstly, we introduce a novel multiple imputation technique called HIMA, which is specifically tailored for high-dimensional neuroimaging data. One of its notable advantages is its efficient computational time, reducing the duration from 2300 hours with classic methods to just 1 hour when dealing with a missing dataset of 500 variables on a single node. Secondly, HIMA aids in expanding brain map coverage by generating accurate multiply imputed results. This is supported by extensive simulation studies that demonstrate reduced bias and dispersion in the imputed data. Benefiting from much-expanded map coverage, HIMA helps to enhance interpretations of imaging results greatly. Lastly, we have developed a user-friendly package that implements HIMA, making it easily accessible and convenient for researchers to utilize this technique in their neuroimaging studies.

The rest of this chapter is organized as follows. In Section 2, we present HIMA model, posterior mode estimation, and algorithms. In Section 3, We assessed the performance of HIMA on two MRI imaging datasets (semi-synthetic and real data) with comparisons to frequently used imputation methods. We conclude with discussions in Section 5.

3.2 Methods

3.2.1 Background

We introduce HIMA on imputing voxel-level missingness in neuroimaging data (e.g., voxel-level hemodynamic response for fMRI[1] data, voxel-level fractional anisotropy for DTI[2] data, ALFF[3] for rs-fMRI[4] among many). Let y_{ij} represent the value of the j-th voxel for the i-th subject, where $i \in [n] := \{1, \ldots, n\}$, and $j \in [p]$. Because most voxel-level neuroimaging variables are continuous variables, we assume $\mathbf{Y} := \{y_{ij}\}_{i\in[n], j\in[p]}$ following a multivariate normal distribution. Without loss of generality, we specify a joint imputation model for $\mathbf{Y}$ based on the regressionAzur et al. (2011); Rubin (2004); Schafer (1997): $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\mathbf{X} \in \mathbf{R}^{n\times r}$, $\boldsymbol{\beta} \in \mathbf{R}^{r\times p}$, and $\boldsymbol{\epsilon}^{n\times p}$ denote the covariates, coefficients, and the error term, respectively.

The p-variate MVN distributions of $\mathbf{Y}$ and $\boldsymbol{\epsilon}$ can be specified as :

$$\mathbf{Y} \sim \mathcal{N}_p(\boldsymbol{\mu} \in \boldsymbol{R}^p, \boldsymbol{\Sigma} \in \boldsymbol{R}^{p\times p}),\ \boldsymbol{\epsilon} \sim \mathcal{N}_p(\mathbf{0}, \boldsymbol{\Sigma}).$$

The observed and missing parts of $\mathbf{Y}$ can be loosely denoted by $\mathbf{Y}_{\text{obs}}$ and $\mathbf{Y}_{\text{mis}}$. Under the commonly used assumption of missing at random (ref neuroimaging and missing data), our goal is to find a set of plausible values to fill in $\mathbf{Y}_{\text{mis}}$ based on the observed data $\mathbf{Y}_{\text{obs}}$ so that the uncertainty of $\mathbf{Y}_{\text{mis}}$ can be preserved. Since $\mathbf{Y}$ follows an MVN distribution, the conditional distribution of its subset vectors $\mathbf{Y}_{\text{mis}}$, given the complement vectors $\mathbf{Y}_{\text{obs}}$, follows an MVN

[1] Functional Magnetic Resonance Imaging
[2] Diffusion Tensor Imaging
[3] Amplitude of low-frequency fluctuation
[4] Resting state fMRI

distribution as well:

$$\mathbf{Y}_{\text{mis}}|\mathbf{Y}_{\text{obs}} \sim \mathcal{N}_p(\boldsymbol{\mu}_{\text{mis|obs}}, \boldsymbol{\Sigma}_{\text{mis|obs}}), \tag{3.1}$$

where

$$\boldsymbol{\mu}_{\text{mis|obs}} = \boldsymbol{\mu}_{\text{mis}} + \boldsymbol{\Sigma}_{\text{mis,obs}}{\boldsymbol{\Sigma}_{\text{obs,obs}}}^{-1}(\mathbf{Y}_{\text{obs}} - \boldsymbol{\mu}_{\text{obs}}), \tag{3.2}$$

$$\boldsymbol{\Sigma}_{\text{mis|obs}} = \boldsymbol{\Sigma}_{\text{mis,mis}} - \boldsymbol{\Sigma}_{\text{mis,obs}}{\boldsymbol{\Sigma}_{\text{obs,obs}}}^{-1}\boldsymbol{\Sigma}_{\text{obs,mis}}. \tag{3.3}$$

Here, $\boldsymbol{\mu} = [\boldsymbol{\mu}_{\text{mis}}, \boldsymbol{\mu}_{\text{obs}}]$ is the partitioned mean vector of two sets of variables, $\boldsymbol{Y}_{\text{mis}}$ and $\boldsymbol{Y}_{\text{obs}}$. Similarly, sub-covariance-matrices are partitioned from $\boldsymbol{\Sigma}$, which captures the information of both voxel-wise and subject-wise relationships, allowing us to jointly utilize different levels of associations for imputation. It is important to recognize that subset vectors $\mathbf{Y}_{\text{mis}}$ and $\mathbf{Y}_{\text{obs}}$ are not of fixed dimension as different subjects may have different missing patterns.

Classical multiple imputation imputed missingness based on statistical characteristics of the data, for example, the distributions of variables in the data set. Bayesian models are commonly used for MI because model parameters ($\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$) and $\mathbf{Y}_{\text{mis}}$ can be iteratively updated. The imputed datasets can be conveniently sampled from the converged posterior distributions (Enders, 2022; Rubin, 2004). However, computational challenges arise for these classic MI methods primarily due to the complexity of $\boldsymbol{\Sigma}$, when missing values are high-dimensional and correlated neuroimaging data.

3.2.2 HIMA model

To address the aforementioned challenges, we propose a relaxed multivariate imputation approach, HIMA, for handling missingness in data with high dimension ($n \ll p$). HIMA is also built upon the Bayesian missing data imputation framework. In Bayesian analysis, *Maximum a Posterior* (MAP) estimation is designed to maximize a conditional probability distribution (i.e., the mode), and sampling the mode of a posterior distribution has been carefully studied in many statistical literatures, with notable examples including Doucet et al. (2002); Gelman et al. (1995); Geman and Geman (1984). As pointed out in Schafer (1997), maximizing $p(\mathbf{\Sigma}|\mathbf{Y})$ is nearly identical to obtaining the ML estimates by maximizing the normal likelihood $L(\mathbf{\Sigma}|\mathbf{Y}) = \Pi_{i=1}^{n}\, p(\mathbf{Y}|\mathbf{\Sigma})$. We incorporate the idea of approximate posterior mode into HIMA. Specifically, instead of randomly drawing $\mathbf{\Sigma}$ from the entire support of the posterior distribution, we draw it from the posterior mode $\mathbf{\Sigma}^{\text{mode}} = \arg\max_{\Sigma_i} p(\mathbf{\Sigma} = \Sigma_i|\mathbf{Y}_{\text{obs}}, \mathbf{Y}_{\text{mis}})$, the value that is most likely to be sampled.

Drawing $\mathbf{\Sigma}^{\text{mode}}$ eliminates the need to invert $\mathbf{\Sigma}_{p\times p}$ with large p and simulate the full posterior distribution at each iteration. This relaxation can significantly alleviate heavy computational burdens from $\mathcal{O}(Cp^3)$ to $\mathcal{O}(Cp)$, where C depends on the number of observations, iterations in MCMC and the number of imputed datasets. Furthermore, avoiding the sampling of large covariance matrices $\mathbf{\Sigma}_{p\times p}$ during each iteration can significantly increase computational stability and trackability, and consequently the accuracy of imputation results. The posterior mode can be estimated by advanced empirical Bayesian methods. We present a tailored method in the following section, which is particularly suitable for data with $n \ll p$.

We now present the proposed HIMA model. On a high level, HIMA aims to iteratively

impute missingness based on Bayesian models as follows:

Step I. Generate parameters $\boldsymbol{\theta}^*$ from the posterior density:

$$\boldsymbol{\theta}^* \sim p(\boldsymbol{\theta}|\mathbf{Y}_{\text{obs}}) = \frac{p(\mathbf{Y}_{\text{obs}}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int p(\mathbf{Y}_{\text{obs}}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}}$$

using Bayes' theorem, where $\pi(\boldsymbol{\theta})$ is the prior density of $\boldsymbol{\theta}$.

Step II. Generate data $\mathbf{Y}_i := \{\mathbf{Y}_{ij}\}_{j\in[p]}$ from the joint imputation model evaluated at $\boldsymbol{\theta}^*$: $\mathbf{Y}_i \sim p(\mathbf{Y}_i|\boldsymbol{\theta}^*)$.

A limitation of the process above, however, is that in many incomplete-data scenarios, simulating the observed-data posterior $p(\boldsymbol{\theta}|\mathbf{Y}_{\text{obs}})$ can be complicated and intractable. Nonetheless, if we augment the observed data $\mathbf{Y}_{\text{obs}}$ with the latent (unobserved) data $\mathbf{Y}_{\text{mis}}$, it becomes easier to simulate the complete-data posterior $p(\boldsymbol{\theta}|\mathbf{Y}_{\text{obs}}, \mathbf{Y}_{\text{mis}})$. Many studies have been conducted to develop data augmentation algorithms, which usually utilize an iterative MCMC procedure to generate imputed values by assuming an underlying distribution (Gelman et al. (2014); Li (1988); Tanner and Wong (1987)). Following these proposed data augmentation algorithms, HIMA is designed to iterate the following two steps until convergence:

The imputation I-step. Given parameters $\boldsymbol{\theta}^{[t-1]}$ at the $(t-1)$-th iteration, generate missing values $\mathbf{Y}_{\text{mis}}^{[t]}$:

$$\mathbf{Y}_{\text{mis}}^{[t]} \sim p(\mathbf{Y}_{\text{mis}}|\mathbf{Y}_{\text{obs}}, \boldsymbol{\theta}^{[t-1]}),$$

which in our case (multivariate normal data) is

$$\mathbf{Y}_{\text{mis}}^{[t]} \sim \mathcal{N}_{\mathbf{Y}_{\text{mis}}|\mathbf{Y}_{\text{obs}}}(\boldsymbol{\mu}_{\text{mis|obs}}^{[t-1]}, \boldsymbol{\Sigma}_{\text{mis|obs}}^{[t-1]}), \tag{3.4}$$

where $\boldsymbol{\mu}_{\text{mis|obs}}^{[t-1]}, \boldsymbol{\Sigma}_{\text{mis|obs}}^{[t-1]}$ are calculated based on $\boldsymbol{\mu}^{[t-1]}, \boldsymbol{\Sigma}^{[t-1]}$ by equations (3.2)(3.3).

The relaxed posterior P-step. Instead of traditional posterior P-step which generates parameters $\boldsymbol{\theta}^{[t]}$ from its posterior $\mathbf{Y}_{\text{mis}}^{[t]}$: $\boldsymbol{\theta}^{[t]} \sim p(\boldsymbol{\theta}|\mathbf{Y}_{\text{obs}}, \mathbf{Y}_{\text{mis}}^{[t]})$. HIMA implements a relaxed P-step by performing the following two sub-steps:

P-step1. Given a complete sample augmented by previous imputed values $\mathbf{Y}_{\text{mis}}^{[t]}$, sample the normal covariance $\boldsymbol{\Sigma}^{[t]}$ from the mode of its posterior density $p(\boldsymbol{\Sigma}|\mathbf{Y}_{\text{obs}}, \mathbf{Y}_{\text{mis}}^{[t]})$, where the mode is obtained by:

$$\boldsymbol{\Sigma}^{[t]} = \arg\max_{\hat{\boldsymbol{\Sigma}}} \; p(\boldsymbol{\Sigma}|\mathbf{Y}_{\text{obs}}, \mathbf{Y}_{\text{mis}}^{[t]}),$$

which in our case is

$$\begin{aligned} \boldsymbol{\Sigma}^{[t]} &= \arg\max_{\hat{\boldsymbol{\Sigma}}} \mathbf{W}^{-1}(\boldsymbol{\Psi} + n\boldsymbol{S}, \nu + n) \\ &= \arg\max_{\hat{\boldsymbol{\Sigma}}} \Big(\frac{|\boldsymbol{\Psi} + n\boldsymbol{S}|^{\frac{\nu+n}{2}}}{2^{\frac{(\nu+n)p}{2}} \Gamma_p\left(\frac{\nu+n}{2}\right)} |\boldsymbol{\Sigma}|^{\frac{-(\nu+n+p+1)}{2}} \exp\Big(-\frac{1}{2} \operatorname{tr}((\boldsymbol{\Psi} + n\boldsymbol{S})\boldsymbol{\Sigma}^{-1})\Big)\Big), \end{aligned} \tag{3.5}$$

where $\boldsymbol{S}$ is the sample covariance $\frac{\sum(\mathbf{Y}_i - \bar{\mathbf{Y}})(\mathbf{Y}_i - \bar{\mathbf{Y}})^{\intercal}}{n}$, and $\mathbf{Y}_i$ is the i-th row from $(\mathbf{Y}_{\text{obs}}, \mathbf{Y}_{\text{mis}}^{[t+1]})$. The posterior (3.5) is derived based on prior $\boldsymbol{\Sigma} \sim \mathbf{W}^{-1}(\boldsymbol{\Psi}, \nu)$.

P-step2. Given augmented data and $\boldsymbol{\Sigma}^{[t]}$, generate normal mean $\boldsymbol{\mu}^{[t]}$:

$$\boldsymbol{\mu}^{[t]} \sim p(\boldsymbol{\mu}|\boldsymbol{\Sigma}^{[t]}, \mathbf{Y}_{\text{obs}}, \mathbf{Y}_{\text{mis}}^{[t]}),$$

which in our case is

$$\boldsymbol{\mu}^{[t]} \sim \mathcal{N}(\bar{\boldsymbol{Y}}, \frac{\boldsymbol{\Sigma}^{[t]}}{n}), \tag{3.6}$$

using a non-informative prior (uniform over the p-dimensional real space)(Schafer, 1997).

When data have n that is not substantially larger than p, or when there are strong relationships among the variables causing certain linear combinations of the columns of $\boldsymbol{Y}$ exhibiting little to no variability, the resulting sample covariance matrix $\boldsymbol{S}$ is sometimes singular or nearly singular. When this happens (which is a common occurrence in brain imaging data), it is not easy to obtain sensible inference about $\boldsymbol{\mu}$. In such situations, it is often suggested to choose a non-informative prior to stabilize the inference process when there is little prior knowledge available about $\boldsymbol{\mu}$ (Box and Tiao, 2011; Schafer and Yucel, 2002).

To summarize, HIMA retains the traditional *imputation I-step* in classical multiple imputation and introduces relaxation to the *posterior P-step*. Both steps are iterated sufficiently long until the sequence $\{(\mathbf{Y}_{\text{mis}}^{[t]}, \boldsymbol{\mu}^{[t]}, \boldsymbol{\Sigma}^{[t]}) : t = 1, 2, \dots\}$ (an MCMC sequence) converges to a stationary distribution $\mathbb{P}(\mathbf{Y}_{\text{mis}}, \boldsymbol{\mu}, \boldsymbol{\Sigma}|\mathbf{Y}_{\text{obs}})$. We have provided theoretical justifications for this procedure in **Appendix 3B**.

3.2.3 Posterior mode estimation

We now present a robust method to estimate the posterior mode of covariance $\mathbf{\Sigma}$ at each iteration in the *relaxed posterior P-step*, which is well-suited for scenarios $n \ll p$. The most natural conjugate class for normal data is the IW family. Following this, we choose IW as the prior distribution for $\mathbf{\Sigma}$. Here, we use an unorthodox representation for IW denoting as $\mathbf{W}^{-1}(\boldsymbol{\zeta}, \lambda)$ for ease of demonstration, where $\boldsymbol{\zeta}$ represents the mean and λ represents a measure of precision depending on the standard degrees of freedom v by $v = \lambda + p + 1$. Under the notations, the prior density is written as

$$\boldsymbol{\zeta} \sim \frac{|\lambda \boldsymbol{z}|^{\frac{\lambda+p+1}{2}}}{2^{\frac{p(\lambda+p+1)}{2}} \pi^{\frac{p(p-1)}{4}} \prod_{j=0}^{p-1} \Gamma\left(\frac{\lambda+j}{2}+1\right)} |\boldsymbol{\zeta}|^{\frac{-(\lambda+2p+2)}{2}} \exp\left\{-\frac{1}{2} \operatorname{tr} \lambda \boldsymbol{z} \boldsymbol{\zeta}^{-1}\right\}.$$

We multiply this prior by the observation probabilities to obtain a joint posterior distribution proportional to

$$|\boldsymbol{\zeta}|^{\frac{-(n+\lambda+2p+2)}{2}} \exp\left\{-\frac{1}{2} \operatorname{tr}(\lambda \boldsymbol{z} + S) \boldsymbol{\zeta}^{-1}\right\},$$

where the mode of this posterior is given by

$$\mathbf{\Sigma}^{\text{mode}} = \frac{\lambda \boldsymbol{z} + \boldsymbol{S}}{n + \lambda + 2p + 2}. \tag{3.7}$$

Here, $\boldsymbol{S}$ is the sample covariance $\sum (\mathbf{Y}_j - \mu)(\mathbf{Y}_j - \mu)^\top$, and λ, $\boldsymbol{z}$ can be estimated using an empirical Bayesian approach proposed in Champion (2003). When λ is large, the prior will swamp the effect of observations. The estimated value of λ tends to be small when the assumed model

is not a good fit for the data. In such cases, the estimation algorithm has no sensible way to proceed (Champion, 2003). However, when the model fits the data well, good estimation can still be achieved even when $n \ll p$. The essential criterion for the optimality of the estimation is to minimize the expected estimation loss by using an estimator $\hat{\boldsymbol{\Sigma}}$ in place of the true parameter $\boldsymbol{\Sigma}_0$, where the estimation loss was defined using the Kullback–Leibler distance:

$$\arg\min_{\hat{\boldsymbol{\Sigma}}} L(\boldsymbol{\Sigma}_0, \hat{\boldsymbol{\Sigma}}) = \int \mathrm{d}\mathbf{Y}\; p(\mathbf{Y} \mid \boldsymbol{\Sigma}_0) \log \frac{p(\mathbf{Y} \mid \boldsymbol{\Sigma}_0)}{p(\mathbf{Y} \mid \hat{\boldsymbol{\Sigma}})}.$$

A great advantage of Champion's method is that, there are no steps in the estimation algorithm that require $n \gg p$. In addition, no steps involve the inversion of large matrices, resulting in enhanced computational efficiency. We provide the asymptotic properties of the posterior mode estimator $\hat{\boldsymbol{\Sigma}}^{\text{mode}}$ in **Appendix 3B**.

3.2.4 Algorithms

In this section, we provide the complete procedures of HIMA in Algorithm 4. In the algorithm, $\mathbf{Y}_0$ denotes the initial dataset before imputation, and $\mathbf{Y}_{(m)}$ denotes the m-th imputed dataset, where $m \in [M]$. The imputed sets $\{\mathbf{Y}_{(m)}\}_{m \in M}$ are identical for the observed data entries, but different in the missing spots. The magnitude of the difference reflects uncertainty around the "truthfulness" of the imputed values. Furthermore, there are a total of T iterations within each $\mathbf{Y}_{(m)}$: by the end of T iterations, the complete imputation is accomplished, resulting in one completely-imputed dataset $\mathbf{Y}_{(m)}$. In other application scenarios, statistical inference can be further made by collectively pooling the estimated results by various pooling methods (e.g., Rubin's rules outlined in Rubin (2004)). Pooling data from $\{\mathbf{Y}_{(m)}\}_{m \in M}$ generally yields improved

inference results.

Algorithm 4 HIMA

procedure ALGORITHM(Input: $\mathbf{Y_0}$, M, T, S)
  for $m = 1, 2, \ldots, M$ imputed datasets **do**
    Impute $\mathbf{Y}_0$ initially[3], assign $\mathbf{Y}$ to the resulted complete dataset
    Initialize $\boldsymbol{\mu}^{[0]} =$ sample mean$(\mathbf{Y})$, $\boldsymbol{\Sigma}^{[0]} =$ sample cov$(\mathbf{Y})$
    for $t = 1, 2, \ldots, T$ iterations **do**
      for $s = 1, 2, \ldots, S$ subjects **do**
        Impute $\mathbf{Y}_{\text{mis}}^{[t+1,s]} | (\mathbf{Y}_{\text{obs}} \sim \mathcal{N}(\boldsymbol{\mu}_{\text{mis|obs}}^{[t,s]}, \boldsymbol{\Sigma}_{\text{mis|obs}}^{[t,s]})$

          where $\boldsymbol{\mu}_{\text{mis|obs}}^{[t,s]} = \boldsymbol{\mu}_{\text{mis}}^{[t,s]} + \boldsymbol{\Sigma}_{\text{mis,obs}}^{[t,s]} \boldsymbol{\Sigma}_{\text{obs,obs}}^{[t,s]\ -1} (\mathbf{Y}_{\text{obs}}^{[t,s]} - \boldsymbol{\mu}_{\text{obs}}^{[t,s]})$

          $\boldsymbol{\Sigma}_{\text{mis|obs}}^{[t,s]} = \boldsymbol{\Sigma}_{\text{mis,mis}}^{[t,s]} - \boldsymbol{\Sigma}_{\text{mis,obs}}^{[t,s]} \boldsymbol{\Sigma}_{\text{obs,obs}}^{[t,s]\ -1} \boldsymbol{\Sigma}_{\text{obs,mis}}^{[t,s]}$

        Estimate posterior mode $\boldsymbol{\Sigma}^{[t]} | (\mathbf{Y}_{\text{obs}}, \mathbf{Y}_{\text{mis|obs}}^{[t,s]}, \boldsymbol{\mu}^{[t]})$
        Sample $\boldsymbol{\mu}^{[t]} \sim p(\boldsymbol{\mu} | \boldsymbol{\Sigma}^{[t]}, \boldsymbol{Y}_{\text{obs}}, \boldsymbol{Y}_{\text{mis}}^{[t,s]})$
      end for
    end for
    Assign the m-th imputed dataset $\mathbf{Y}_{(m)} = (\mathbf{Y}_{\text{obs}}, \mathbf{Y}_{\text{mis}}^{[t,s]})$
  end for
  return $\{\mathbf{Y}_{(m)}\}_{m \in [M]}$
end procedure

Simulations have indicated that HIMA can perform well with a small number of imputations M (e.g., $5 - 20$). The number of iterations T can be specified by researchers. According to our practice, a low number of cycles (e.g., $20 - 30$) is typically sufficient for the convergence of parameters and imputed values. Previous imputations $\bar{\mathbf{Y}}_{\text{mis|obs}}^{[t-1]}$ only indirectly enter $\bar{\mathbf{Y}}_{\text{mis|obs}}^{[t]}$ through its relation with observed voxels $\mathbf{Y}_{\text{obs}}$, which leads to fast convergence.

3.3 Data example

A real MRI dataset from a schizophrenia study is used to assess the accuracy and efficiency of HIMA compared to existing imputation methods. The study focus was on cross-sectional

[4]Using feasible method (mean imputation, median imputation, hot deck imputation, etc.)

neurovascular water exchange (Kw) data collected from 58 subjects (age, M = 37.8, SD = 14.0; sex, 37 M: 21 F) with schizophrenia spectrum disorder. HIMA is applied to two different versions of the data: (i) semi-synthetic data that contained partially known missing values (i.e., original unknown missing entries plus known missingness that are manually inserted). (ii) real data that contained completely unknown missing values.

The subjects were recruited from mental health clinics and media advertisements in Baltimore, MD. Imaging data were collected based on a diffusion-prepared arterial spin labeling protocol from 2019-2022 using a Siemens 3T PRISMA scanner with a 64-channel head coil. A comprehensive explanation of imaging acquisition and processing procedures can be found in **Appendix 3A**. This schizophrenia study aims to investigate the association between neural-capillary water exchange function and schizophrenia spectrum disorder, with a specific focus on 99 brain regions of interest obtained using the International Consortium for Brain Mapping (ICBM) brain template.

3.3.1 Semi-synthetic data analysis

The primary measure of this schizophrenia study was the whole-brain average Kw values. We intend to apply our method to impute missing values within each of the 99 regions individually because within-region Kw values are typically more homogeneous and strongly associated. Accordingly, the data structure to be imputed is $\{y_{ij}\}_r$, where $i \in [58]$ subjects, $r \in [99]$ regions, and the number of voxels $j \in [p_r]$ depends on region r. The overall missing rate of data $\{\{y_{ij}\}_r : r \in [99]\}$ is 14.60%. On top of this, we construct a semi-synthetic dataset by manually removing data in random (i, j), i.e., random voxel locations of random subjects. Specifically, we randomly remove t data entries from each voxel vector, where t can be any number from 1 to 8.

This results in, on average, 4.5 extra missing spots (out of 58) in each voxel vector The missing rate of the semi-synthetic dataset is now 20.34%, compared to the original rate of 14.60%.

A few processing steps on the semi-synthetic dataset are as follows. First, we drop out voxel vectors that have remarkably high missing rates (we set 40% as the threshold) to ensure robust performance. Experiments show that in this dataset, HIMA may become unstable and less accurate after the voxel-vector-level threshold is set higher than 40%. On average, only 3.17% voxels exceed the threshold within each region (e.g., for a region with 800 voxels, only 25 voxels need to be dropped out). Next, we perform kernel smoothing on subjects and align their smoothed probability density estimate to take into account subject-level random effect. Lastly, we perform HIMA to impute the missing data on each post-processed $\mathbf{Y}_r^* = \{y_{ij}^*\}_r$, where we set $M = 15$ (number of imputed datasets) and $T = 20$ (number of iterations).

To evaluate the imputation performance, we first define an $n \times p$ indicator matrix $\mathbf{R}$ for a data matrix $\mathbf{Y} = \{y_{ij}\}_{i\in[n],j\in[p]}$, where each element $r_{ij} = 1$ if y_{ij} is observed; $r_{ij} = 0$ otherwise. Using $\mathbf{R}$, we assess imputation accuracy by the following metrics:

(i) Weighted mean absolute error (wMAE):

$$\frac{\sum_{i\in[n]}\sum_{j\in[p]} I(r_{ij}=0) \times |y_{ij}^{\text{imputed}} - y_{ij}^{\text{true}}| \,/\sqrt{\text{Var}(Y_{\cdot j})}}{\sum_{i\in[n]}\sum_{j\in[p]} I(r_{ij}=0)}$$

for a single imputed dataset. We then compute the mean and standard deviation of wMAE across all M imputed datasets.

(ii) Weighted mean square error (wMSE):

$$\frac{\sum_{i\in[n]}\sum_{j\in[p]} I(r_{ij}=0) \times (y_{ij}^{\text{imputed}} - y_{ij}^{\text{true}})^2 /\sqrt{\text{Var}(Y_{\cdot j})}}{\sum_{i\in[n]}\sum_{j\in[p]} I(r_{ij}=0)}$$

for a single imputed dataset. Again, we collect the mean and standard deviation of wMSE across all M imputed datasets.

(iii) Weighted mean bias error (wMBE):

$$\frac{\sum_{i\in[n]}\sum_{j\in[p]} I(r_{ij}=0)\times(\mathbb{E}[y_{ij}{}^{\text{imputed}}]-y_{ij}^{\text{true}})/\sqrt{\text{Var}(Y_{\cdot j})}}{\sum_{i\in[n]}\sum_{j\in[p]} I(r_{ij}=0)}$$

across all M imputed datasets, where $\mathbb{E}[y_{ij}{}^{\text{imputed}}]$ can be estimated by $\frac{\sum_{m\in[M]} y_{ij}^{(m)\text{imputed}}}{M}$.

We select three regions for result demonstration: right insular cortex (Ins), right caudate nucleus (Caud), and right hippocampus (Hippo), which are brain areas frequently associated with information processing in schizophrenia (Harrison, 2004; Mueller et al., 2015). Based on ICBM, there are respectively $1196, 553,$ and 622 voxels in the right Ins, right Caud, and right Hippo. We applied HIMA on the post-processed data $\mathbf{Y}^*_{\textbf{Ins}}$, $\mathbf{Y}^*_{\textbf{Caud}}$, $\mathbf{Y}^*_{\textbf{Hippo}}$ and compared the imputation performance with frequently used approaches in both simple imputation (mean-substitution imputation) and multiple imputation (MICE). As mentioned previously, almost all existing multiple imputation methods and toolkits for brain imaging data were based on MICE. We provide the imputation error measure and computational time of these three methods in Table 1. In addition, Figure 3.4 shows the trace plots of estimates over iterations.

Based on Table 1 and Figure 3.4, we evaluate the imputation performance from the following three aspects:

1. **Accuracy:** Based on various error metrics (wMAE, wMSE, and wMBE), HIMA shows the lowest imputation errors among all three selected brain regions, followed by Mean imputation and MICE. Furthermore, the imputed results generated by HIMA also exhibit the

Table 1. Imputation performance of kw data in schizophrenia study (M=15, T=20)

	wMAE	wMSE	wMBE	Time	Compiler
			Region : *Right insular*		
Mean Imputation	0.82(0)	35.07(0)	4.6×10^{-3}	1s	MATLAB 2021, 16 GB memory, single cluster node
MICE	1.22(0.03)	96.62(2.56)	-0.24	≈756hrs (≈ 15120 hrs if using single node)	R 2022, 128 GB memory, 20 cluster nodes
HIMA (Proposed)	$0.31(3.33\times10^{-5})$	$7.13(1.20\times10^{-3})$	5.23×10^{-4}	4.15h	MATLAB 2021, 16 GB memory, single cluster node
			Region : *Right caudate nucleus*		
Mean Imputation	0.80(0)	35.58(0)	-8.19×10^{-3}	1s	MATLAB 2021, 16 GB memory, single cluster node
MICE	$1.35(9.75\times10^{-3})$	98.91(2.74)	-0.32	≈210 hrs (≈4200 hrs if using single node)	R 2022, 128 GB memory, 20 cluster nodes
HIMA (Proposed)	$0.27(5.84\times10^{-5})$	$6.27(2.40\times10^{-3})$	-7.11×10^{-3}	1.32h	MATLAB 2021, 16 GB memory, single cluster node
			Region: *Right hippocampus*		
Mean Imputation	0.71(0)	34.54(0)	-0.01	1s	MATLAB 2021, 16 GB memory, single cluster node
MICE	1.16(0.02)	90.20(2.31)	-0.08	≈115 hrs (≈2300 hrs if using single node)	R 2022, 128 GB memory, 20 cluster nodes
HIMA (Proposed)	$0.21(3.40\times10^{-5})$	$4.41(1.22\times10^{-3})$	-4.31×10^{-3}	0.96h	MATLAB 2021, 16 GB memory, single cluster node

***Region info**:	Right insular	# voxels: 1134	Missing rate: 8.31%	Kw value: 108.29(37.22)
	Right caudate nucleus	# voxels: 549	Missing rate: 8.30%	Kw value: 112.69(38.02)
	Right hippocampus	# voxels: 500	Missing rate: 17.83%	Kw value: 98.03(42.36)

Figure 3.3:

lowest variances, indicating high consistency and stability of imputed results. In summary, HIMA demonstrates the highest imputation accuracy, as evidenced by its reduced errors and dispersion.

2. **Convergence:** We created trace plots to visualize the convergence of the estimated mean of Kw values against iteration numbers. We randomly selected two voxels (voxels $\#120$ and $\#305$) in the right hippocampus to examine the convergence performance. We observed that both methods had reached a stable posterior distribution after a few iterations, indicating quick convergence to stationarity. With HIMA, voxels converged to stationary estimates

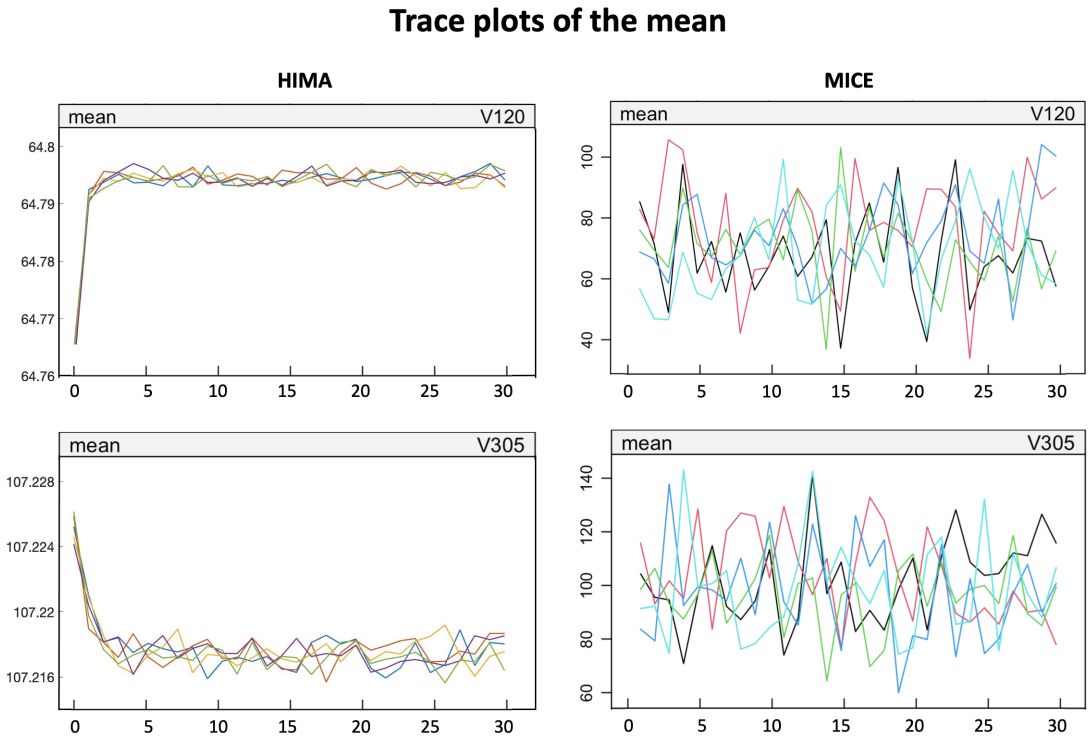


Figure 3.4: *Trace plots of the means for two randomly selected incomplete voxels #120 and #305 in the right hippocampus against the number of iterations. The means were estimated iteratively by methods HIMA and MICE, both of which show trends of convergence to stationary, while HIMA converged faster and had reduced dispersion of estimates. Besides, all plots were free of any abnormal trends.*

with smaller variations, suggesting a higher level of stationarity. Additionally, neither method produced any discernible trends, suggesting sufficient randomness in the estimates across iterations.

3. **Efficiency:** We conducted an evaluation of the computational efficiency of these imputation methods. Mean imputation spends the shortest running time due to its simple operation, without the need to draw and update multiple values during each iteration. However, its imputation accuracy is relatively low. On the other hand, HIMA employs the principle of multiple imputation and demonstrates significantly improved computational efficiency compared to MICE, while still reducing imputation errors. For regions with varying sizes and missing rates, HIMA reduces the time required from up to 10^4 hours (using the popular

MICE method) to just a few hours, indicating a significant advancement in the computational competency for imputing ultra-high dimensional imaging data.

3.3.2 Real data analysis

In this real data analysis, we apply HIMA on each of the 99 distinct regions without manually inserting any missingness this time. Hence, the dataset $\{y_{ij}\}_{r\in[99]}$ to be imputed is the original MRI data, where its missing entries were mainly caused by image acquisition limitations and susceptibility artifacts. The overall missing rate of the data is 14.6%. We first applied the same preprocessing procedures on each $\{y_{ij}\}_r$: (i) exclude voxels with missing rates higher than 40% to ensure HIMA's stability and accuracy (on average, less than 3% voxels in this real dataset were excluded for each region); (ii) align kernel-smoothed probability density estimate of each subject to eliminate subject-level random effect during imputation. We next applied HIMA on each post-processed data $\{\mathbf{Y}_r^*, r \in [99]\}$ with $T = 20$ iterations and $M = 50$ imputations. We increased the number of imputations this time since true information about missing Kw was not available, and we wanted to assess the stability and accuracy of imputed values more carefully.

We used the same three schizophrenia-associated regions (right Ins, right Caud, and right Hippo) to evaluate the imputation performance. The error metrics proposed previously are not applicable here since we don't know the true value of missingness. Instead, we evaluated the results by inspecting the distributions of the original values and the imputed values for voxels within each region. In Figure 3.5(a), we computed the voxel-level Kw mean of 50 imputed datasets $\bar{\mathbf{Y}}_r$ and plotted $\bar{\mathbf{Y}}_r$ against the original data for two randomly selected voxels from region r. We see that the shape of the red points (voxels from $\bar{\mathbf{Y}}_r$) matches the shape of the purple ones (voxels

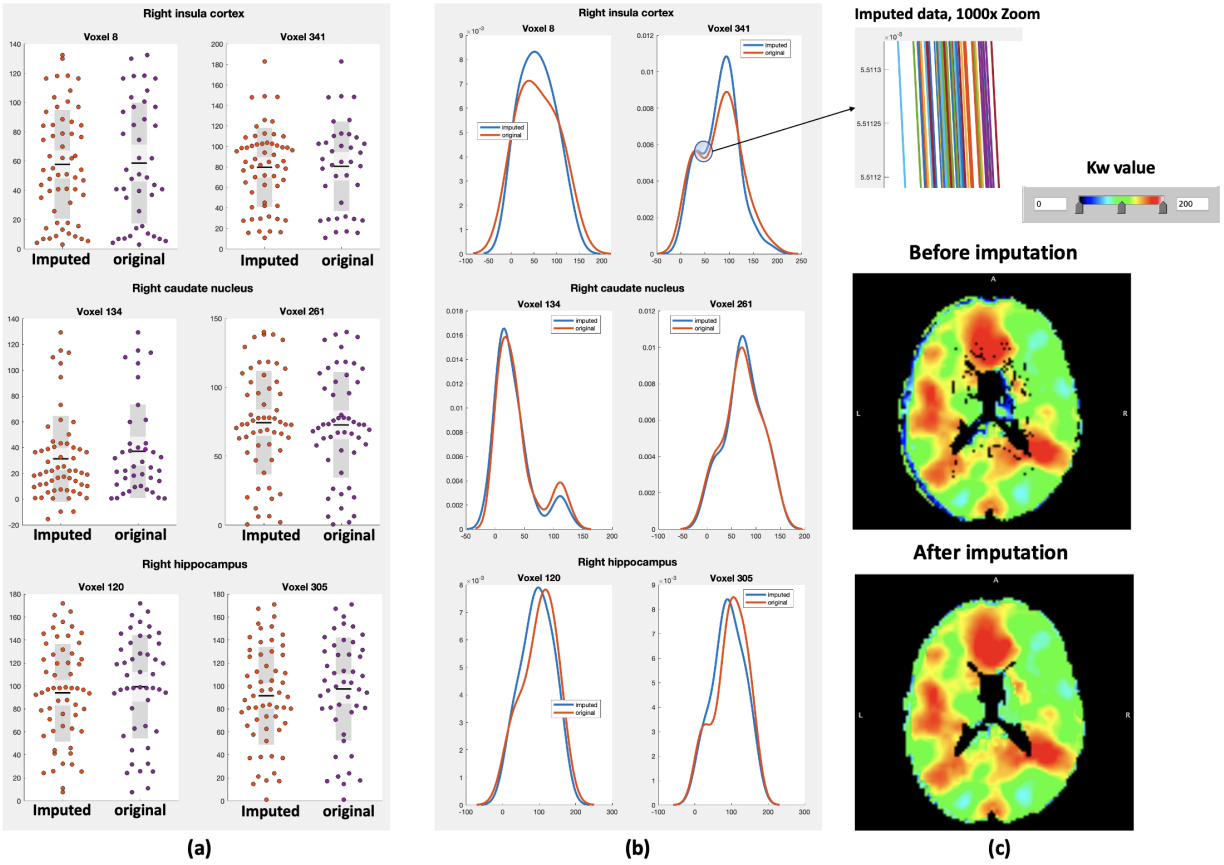


Figure 3.5: *(a) The shape of the red points (voxels from imputed data) matches the shape of the purple ones (voxels from original data). The matching shape indicates the plausibility of the imputed values. (b) The probability density estimate of the original data is similar to the density estimates of all 50 imputed data. The density estimates of 50 imputed data are so close to each other that they appear to be emerging as one single blue curve. On the right side, we showed a 1000-times-closer zoom-in version of the imputed values for the* 341*-th voxel in the right insula cortex. (c) shows the heatmaps of the whole-brain Kw values from a randomly selected subject before and after imputation using HIMA.*

from original data). The matching shape indicates the plausibility of the imputed values. In Figure 3.5(b), we plotted the probability density estimate of the original data against that of all 50 imputed data for the same two voxels from region r. The density estimates of 50 imputed data are so close to each other that they appear to be emerging as one single blue curve on the graph. On the right side, we showed a 1000-times-closer zoom-in version of the imputed values for the 341-th voxel in the right insula cortex. We observed that the distributions of these voxels from all 50 imputed datasets are similar to those from the original dataset. Again, the observation met our

expectations, and no anomalies of outliers were evident.

3.4 Discussion

Reliable analyses of brain imaging studies are often compromised by the common issue of missing data. To address this problem, we introduce a new method called HIMA, which is a multiple imputation approach designed specifically for high-dimensional data under the assumption of MAR. HIMA has the following key features:

1. Missingness uncertainty preserved: A common misconception of many missing data methods (e.g., mean/mode substitutions, linear/spline interpolation) is the assumption that imputed values should represent "real" values. HIMA, on the other hand, acknowledges the uncertainty around the "truthfulness" of the imputed values and incorporates it into the imputed values by drawing a set of plausible values from joint imputation models.

2. Accuracy improved: HIMA adapted the *posterior P-step* by drawing approximate posterior mode, which is the most probable draw from posterior distribution and was shown to have good asymptotic properties. This adaption avoids the need to sample large covariance matrix $\mathbf{\Sigma}_{p\times p}$ (e.g. $p = 1000$) at each iteration which increases computational stability and imputation accuracy. Simulation results from both semi-synthetic and real data demonstrate that HIMA produces robust estimates of the original data, as indicated by significantly smaller weighted errors and dispersion.

3. Computational efficiency boosted: Compared to the widely-used MICE package, HIMA offers significant improvements in computational efficiency. Without compromising accuracy,

HIMA lowers the computational complexity from $\mathcal{O}(Cp^3)$ to $\mathcal{O}(Cp)$ For imaging data with large p typically, this cab lead to a remarkable improvement in imputation efficiency. For instance, in regions with hundreds of voxels, HIMA reduces the running time from thousands of hours required by MICE to a few hours.

There are several considerations to be made before performing HIMA regarding missing proportions, number of imputations, and number of iterations that can affect imputation quality. Simulations have indicated that HIMA can perform well with up to 40% missing observations. However, the more missing information there is, the more likely the imputation procedure may encounter estimation issues and the less likely that the MAR assumption is met unless it was planned missing. Simulations have also shown that HIMA generally produces low-biased imputation with low variance, so a small number of imputations M (e.g., 5 to 20) is fair enough. Also, a moderate number of iterations (e.g., 15 to 30) is typically sufficient for HIMA to reach an appropriate stable posterior distribution, indicating a fast convergence to stationarity.

In practice, nearest neighbor gaussian processes (NNGP) based models are often leveraged in neighbor-based imputation. NNGP, however, may not be suitable for brain imaging data because voxels in brain are organized in complex structures blocked by ridges and sulci, and the voxel-pair correlation is not directly determined by their spatial proximity. We provide a plot of voxel-pair correlations against voxel-pair spatial distance in **Appendix** 3C, where the plotted points are approximately circular rather than linear. For this reason, we did not incorporate NNGP in our method. Another point worth pointing out is that, a common limitation of many multiple imputation approaches is that fitting a series of conditional distributions may not be consistent with proper joint distribution. While some initial studies suggest that this may not pose a big issue

in practical circumstances (Brand (1999),Schafer and Yucel (2002)), further research is needed to fully understand these implications.

Multivariate missing data is still an active area of statistical research, and many scholars have made significant efforts to advance imputation techniques. We hope that HIMA can serve as a practical toolkit for researchers to address the very common missingness problem in brain imaging data, especially in high-dimensional space, and thereby enhance interpretations of imaging inference results. The sample codes for HIMA are available at https://github.com/TongLu-bit/HighDim-MultipleImputation-HIMA. In addition to brain imaging data, HIMA may provide a promising solution for other data with high dimensional features, such as data in healthcare, finance, and genome science.

Chapter 4: Multi-level network association analysis (MOAT)

4.1 Introduction

Neuroimaging data plays a fundamental role in deciphering the operations of the human brain, the most complex organ. These data come in various modalities, such as magnetic resonance imaging (MRI), Diffusion Tensor Imaging (DTI), and functional MRI (fMRI), each revealing distinct aspects of the brain's structure and function. For instance, MRI yields high-resolution images of the brain's structure, offering valuable physical information such as size, shape, cortical thickness and more; DTI gauges the integrity of white matter micro-structures by calculating fractional anisotropy; functional MRI (fMRI) captures dynamic blood flow changes in various brain regions to measure localized neural activity and functional connection.

In statistical analysis, neuroimaging data are commonly represented in two structures: vectors (e.g., a list of region-wise cortical thickness measures) and association matrices (e.g., functional connectivity strength stored in an adjacency matrix) (Bullmore and Sporns, 2009; Derado et al., 2010; Fornito et al., 2016; Penny et al., 2011; Wig et al., 2014). The study of brain structural and functional imaging data is the key to gaining insights into the neurophysiological and neuropathological underpinnings of human behavior and cognition. Instead of separately studying brain *structural imaging (SI)* and *functional connectivity (FC)* data, exploring their intricate interplay could significantly deepen our understanding of the brain, including its development and

aging (Bowman et al., 2012; Drevets et al., 2008; Kemmer et al., 2018; Smith et al., 2004). For example, brain regions connected by white matter tracts with higher fractional anisotropy (FA) are more likely to demonstrate strong FCs. These strong connections can consequently influence cognitive processes such as attention, memory, and decision-making.

There exists little work on joint analysis of multi-modal neuroimaging data despite its significant importance, possibly due to the challenge presented by ultra-high dimensionality and intertwined data structures. In conventional brain connectome studies, researchers frequently gather 10^5 FC measures across hundreds of brain regions and up to 10^4 SI measures, resulting in billions (10^9) of FC-SI pairs. This not only creates significant computational demands but also poses challenges for multiple-testing correction. Traditional correction methods like the false discovery rate (FDR) and family-wise error rate (FWER) often yield almost no supra-threshold FC-SI pairs, as demonstrated by extensive simulation studies. Moreover, FC and SI display data structures indicative of the connectomic network space and spatial dependence respectively. A joint FC-SI analysis requires these intertwined data structures to be incorporated into statistical modeling in order to generate biologically plausible and interpretable results. Specifically, our goal of this work is to identify an array of SI variables that intrinsically influences a group of FCs within a brain connectome sub-network, rather than those randomly distributed across the whole-brain connectome (referred to as a *systematic* pattern of associations). These challenges underscore the necessity of developing a *joint analysis method* to address the complexity of multi-modal neuroimaging data.

Recently, advanced statistical methods have been developed to jointly model two sets of neuroimaging features by leveraging techniques including regularization, low rank, and projection models (Kong et al., 2019; Li et al., 2012b; Wang et al., 2011; Zhu et al., 2014). Many of

these methods have been successfully applied to multi-modal imaging data analysis and yielded interesting findings (Ball et al., 2017; Hayden et al., 2006; Wehrle et al., 2020; Zhang et al., 2022). These statistical methods can be broadly classified into two categories. The first category uses regularization-based methods (Wang et al., 2020; Zhou and Li, 2014; Zhu et al., 2017), which select a parsimonious set of FC-SI associations. A major limitation of these methods is that the sparsely selected associations fail to take into account the systematic impacts of SIs on FC networks. The second category employs dimensional reduction strategies, such as principal component analysis (PCA) (Chachlakis et al., 2019; Hotelling, 1933; Jolliffe and Cadima, 2016), which first projects both FCs and SIs into a handful of top principal components and then performs regression analysis on these selected components. However, as an unsupervised dimension reduction technique, PCA-based analysis often extracts fewer associated principal components of outcomes and predictors, thereby missing the truly associated FC-SI pairs. Sparse canonical correlation analysis (sCCA) methods can be considered as an integration of these two categories and have been widely used in neuroimaging studies (Lin et al., 2013; Uurtio et al., 2019; Witten et al., 2009). Yet, sCCA methods usually focus on vector-to-vector association analysis, which may also overlook the systematic vector-to-network association patterns that are of particular interest in in this work (i.e., the associations between the SI vector and FC sub-connectome represented as a matrix). To bridge the methodological gap in modeling vector-to-matrix associations and incorporating latent network structures, we propose a new **m**ulti-level netw**o**rk associ**a**tion me**t**hod (**MOAT**) to systematically investigate the FC-SI association patterns.

Figure 4.1 presents an overview of the MOAT method, which is constructed based on a multi-level graph for structural-functional neuroimaging data. The first level is a bipartite graph that depicts the association patterns between the SI vector as predictors and vectorized

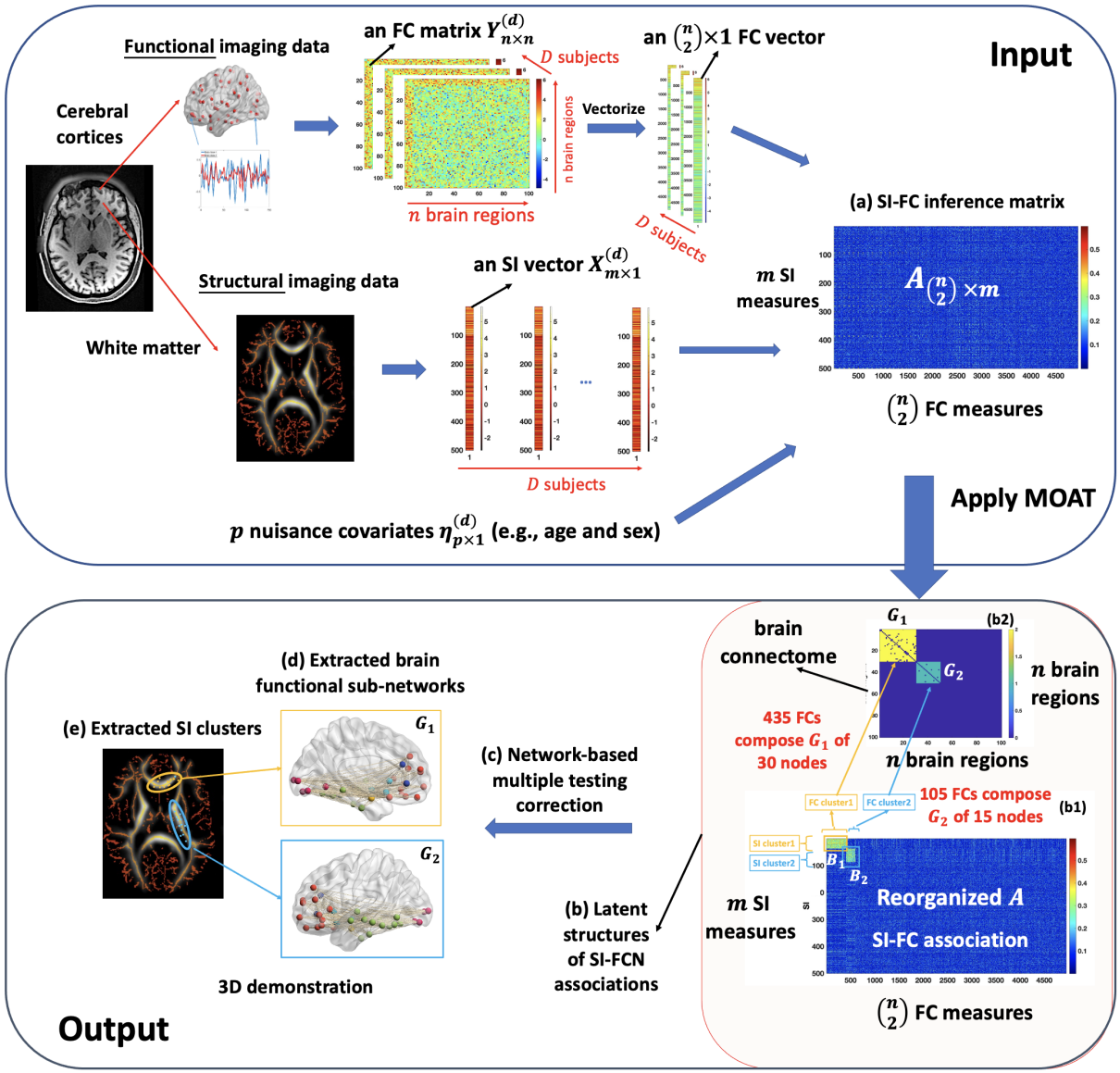


Figure 4.1: *The detection pipeline of systematic FC-SI association patterns by MOAT. (a) shows a heatmap of an association matrix* $\boldsymbol{A}$*, with each element quantifying the association between SI and FC measures (e.g., correlation, test statistics,* $-log(p)$ *value). A hotter point indicates a stronger FC-SI association. (b) shows the systematic FC-SI association patterns identified by MOAT, where the first level (b1) depicts the revealed association patterns between SIs and FCs, where strong associations are mostly concentrated in two dense sub-networks; the second level (b2) depicts the organized structure of a graph formed by the selected FC clusters in (b1). In step (c), our proposed network-based inference test is performed on each FC-SI sub-network identified by MOAT. (d) shows a 3D visualization of the significant FC-SI network-level associations that pass the inference test.*

FC outcomes, adjusted for other confounding covariates (Figure 4.1(a)). Meanwhile, the second level is a complete unipartite graph that reconstructs the vectorized FCs back to a whole-brain connectome network. This multi-level structure enables the identification of subsets of SIs that

systematically impact FC sub-networks. We have further developed computationally efficient algorithms to extract the multi-level sub-networks from the full graphs and have proposed a tailored network-based inference frame to first globally test the existence of the sub-network structures and then individually test each sub-network with multiple corrections based on permutation tests. Our method is also compatible with the existing methods aforementioned (e.g., PCA, CCA). For example, applying CCA to FCs and SIs in an extracted multi-level sub-network provides an estimate of association in the context of multiple regressions.

The contributions of this article are three-fold. First, we introduce MOAT, a novel method that can handle matrix-variate outcomes and vector-variate predictors. Compared to the existing models for multivariate outcomes and multivariate predictors (Lu et al., 2023; Mihalik et al., 2022; Wu et al., 2021; Zhuang et al., 2017), MOAT can further account for the network structure within the matrix outcomes and between the outcome-predictor association patterns. MOAT naturally prohibits most false positive associations because these associations are more likely distributed sparsely rather than gathered in organized sub-networks. Secondly, we develop new algorithms to extract those multi-level sub-networks. The computational load is relatively low because we have developed a tailored greedy peeling algorithm with multilinear complexity, making our approach compatible with the commonly used permutation tests that are often computationally intensive. Lastly, we proposed a novel inference framework at both the global level and individual network level. For the individual level inference, we utilize novel test statistics derived based on the multi-level dense subgraph properties in terms of size and density, which simultaneously improves sensitivity and specificity by leveraging graph combinatorial theories.

The rest of this chapter is organized as follows. In Section 2, we formally define the multi-level network structure and present how MOAT works in network extraction with the

network-based inference method. In Section 3, we perform extensive simulation analyses for model validation and comparison. In Section 4, we apply MOAT to a real structure-function neuroimaging dataset from UK Biobank with 4,242 participants to systematically investigate the FC-SI associations. We conclude with discussions in Section 5.

4.2 Our method

4.2.1 Data structure and problem set up

We observe structural-functional neuroimaging data from D independent subjects. For each subject $d \in \{1, \ldots, D\}$, three sets of measurements are collected:

(i) Independent covariates: a vector of m structural imaging (SI) measures $\boldsymbol{X}^{(d)} = \left(x_1^{(d)}, \ldots, x_m^{(d)}\right)^T$. This vector characterizes anatomical structures of the brain, for example, white matter microstructure integrity, measured by fractional anisotropy from DTI based on the ICBM DTI-81 Atlas (Mori et al., 2008), and region-wise cortical thickness obtained from MRI based on FreeSurfer (Tustison et al., 2014).

(ii) Outcome variables: an adjacency matrix storing pair-wise FC measures $\boldsymbol{Y}_{n\times n}^{(d)}$ between n brain regions. Each element $\left\{y_{ij}^{(d)}\right\}_{i,j<n}$ in $\boldsymbol{Y}^{(d)}$ represents the FC strength between brain regions i and j, obtained from functional imaging data such as resting state fMRI. It is commonly assumed that the FC brain regions are identical and invariant across participants because a common functional brain atlas (e.g., the Brainetcome Atlas (Fan et al., 2016)) is often applied to spatially normalized imaging data. We consider FC measures as random variables due to their inherent variability across different clinical states and conditions.

Moreover, neurologically, it is widely acknowledged that brain structure determines neural functions (Bai et al., 2009; Buckner et al., 2008; Honey et al., 2010).

(iii) Confounding variables: a vector of p confounding covariates $\boldsymbol{\eta}^{(d)} = \left(\eta_1^{(d)}, \ldots, \eta_p^{(d)}\right)^T$. Brain functional connectome is a complex mechanism influenced by multiple factors in addition to the brain structure, including age, sex, genetics, environment, etc. We include them as confounding variables in our model.

To examine the biological association between FCs and SIs, consider the following regression model:

$$y_{ij}^{(d)} = \theta_{(ij),k} + \beta_{(ij),k} x_k^{(d)} + \sum_{p=1}^{P} \alpha_{(ij),k}^{p} \eta_p^{(d)} + \epsilon_{(ij),k}^{(d)}, \tag{4.1}$$

where $\theta_{(ij),k}$ is an intercept term, $\beta_{(ij),k}$ is the coefficient of a localized SI measure $x_k^{(d)}$, $\alpha_{(ij),k}^{p}$ is the coefficient for the confounding covariate $\eta_p^{(d)}$, and $\epsilon_{(ij),k}^{(d)}$ is an error term. In Equation 4.1, we focus on a single SI measure each time in the sense of variable screening because it is challenging to simultaneously handle billions of associations between FCs and SIs in one model in order to accurately identify truly associated FC-SI pairs. (Marek et al., 2022; Mbatchou et al., 2021; Woo et al., 2014). This strategy is commonly used in large-scale imaging and genetics data analysis, followed by further analysis tools to reveal organized association patterns (e.g., in network structures) (Chen et al., 2023; Schaid et al., 2018; Zalesky et al., 2010).

The whole-brain FC-SI association can be summarized in a three-dimensional tensor $\tilde{\boldsymbol{\beta}} = \{\beta_{(ij),k}\}_{\forall ijk} \in \mathbb{R}^{n \times n \times m}$. Alternatively, we can represent $\tilde{\boldsymbol{\beta}}$ using a two-dimensional matrix $\boldsymbol{\beta} \in \mathbb{R}^{\binom{n}{2} \times m}$ by vectorizing the FC matrix $\boldsymbol{Y}_{n \times n}$ into an FC vector $\boldsymbol{Y}_{1 \times \binom{n}{2}}$. Each element

$\beta_{(ij),k} \in \boldsymbol{\beta}$ indicates the strength of association between the functional connection Y_{ij} and the SI measure X_k. In practice, only a small fraction of $\beta_{(ij),k} \neq 0$, reflecting that FCs are sparsely influenced by SIs. However, due to the ultra-high dimensionality of $\boldsymbol{\beta}$, it is challenging to identify $\beta_{(ij),k} \neq 0$ with high sensitivity and rigorously controlled false positive findings. Furthermore, the extraction of latent yet systematic association patterns of between FC and SI (e.g., a set of white matter tracks intensely affecting an FC sub-network via altering the density of neurotransmitters) presents an even more formidable task. To address these challenges, we resort to multi-level graph models.

4.2.2 Multi-level graph structure for $\{\beta_{(ij),k}\}$

We first present the two-dimensional FC-SI association matrix $\boldsymbol{\beta} \in \mathbb{R}^{\binom{n}{2} \times m}$ using a bipartite graph $B = (S, F; H)$ (i.e., *level 1*), where S represents the node set of SI measures with size $|S| = m$, F denotes the node set of FC measures with size $|F| = \binom{n}{2}$, and H denotes the edge set between SIs and FCs with size $|H| = \binom{n}{2} \times m$. An edge $h_{(ij),k} \in H$ corresponds to the bridge between node $s_k \in S$ (the k-th SI) and $f_{i,j} \in F$ (the FC strength between the i-th and j-th brain regions). The left panel in Figure 4.2 shows an example of a bipartite graph B. We consider the node sets of SIs and FCs as two disjoint sets because we are exclusively interested in the associations between FCs and SIs.

Given that the node set F of FC measures represents the whole-brain functional connectome, each node $f_{(i,j)} \in F$ naturally unfolds to an edge in a brain connectome graph G (i.e., *level 2*). Let $G = \{V; F\}$ denote a unipartite graph, where V is the node set of brain regions with size $|V| = n$ and F is the edge set of interregional connections with $|F| = \binom{n}{2}$. The locations of all nodes in V

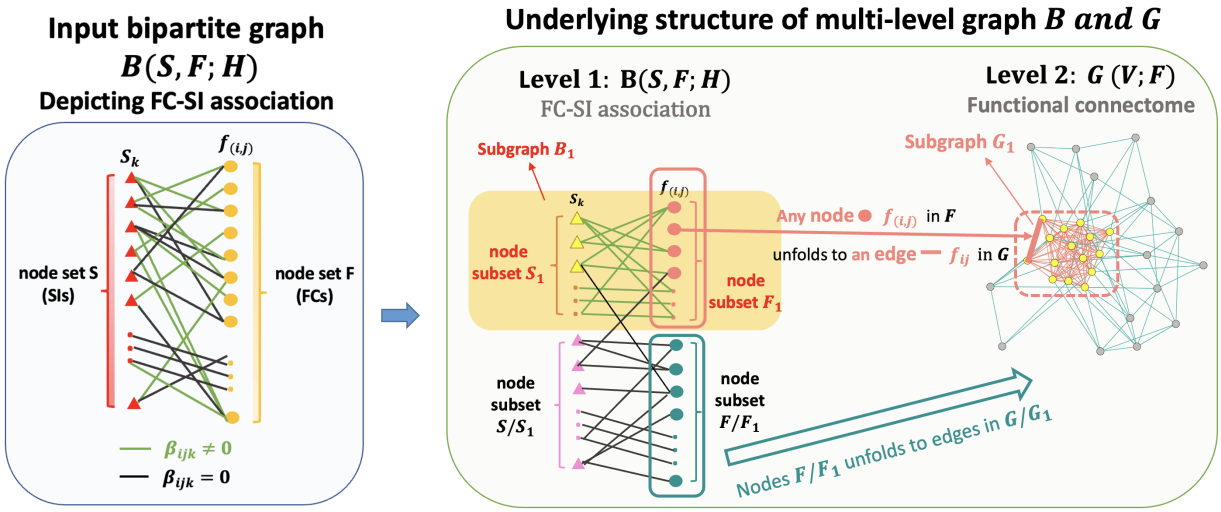


Figure 4.2: *An illustration of a multi-level graph B and G with a FC-SI associated sub-network B_1. Specifically, $B = (S, F; H)$ is an input bipartite graph consisting of two node sets S and F, representing the SI and FC measures respectively. The edge set H connecting S and F can be binary or weighted, where the weight of each edge $h_{(ij),k} \in H$ is determined by the regression coefficient $\boldsymbol{\beta} = \{\beta_{(ij),k}\}$ quantifying the association between y_{ij} (FC strength between brain regions i and j) and x_k ($k-$th SI measure). The right panel shows an example of the multi-level sub-community structures underlying B and G: within B, there exists a sub-network $B_1 = (S_1, F_1; H_1) \subset B$ which contains SI-related edges $\{h_{(ij),k} | i, j, k \in B_1\}$ with a much higher density than the rest $\{h_{(ij),k} | i, j, k \notin B_1\}$. Given that the node set F represent the whole-brain functional connectome, each node $f_{(i,j)} \in F$ naturally unfolds to an edge in a brain connectome graph $G = \{V; F\}$, where V is the node set of brain regions, F is the edge set of interregional connections, and the weight of each edge f_{ij} is also determined by $\boldsymbol{\beta}$. F_c in the sub-network B_c are not isolated edges in brain connectome, instead they constitute an organized sub-graph $G_1 \in G$.*

are pre-specified by a known brain atlas shared across subjects, which consequently determines the spatial location of each edge $f_{ij} \in F$. In our application, an edge f_{ij} conceptually corresponds to the functional connection between a region pair.

We use the proposed multi-level graph model B and G to collectively characterize all potential associations between FC-SI pairs and the network patterns underlying the whole-brain functional connectome space influenced by SIs. In the model, we utilize graph weights to quantify the strength of associations. For example, for each edge $h_{(ij),k} \in H$, the weight $w(h_{(ij),k})$ reflects the association strength of $\beta_{(ij),k}$ between FC Y_{ij} and SI X_k. Without loss of generality, we

consider the following binary weights for all edges in B and G:

$$\begin{cases} w(h_{(ij),k}) = 1 \text{ in } B, w(f_{ij}) = 1 \text{ in } G, & \text{if } \beta_{(ij),k} \neq 0; \\ w(h_{(ij),k}) = 0 \text{ in } B, w(f_{ij}) = 0 \text{ in } G, & \text{if } \beta_{(ij),k} = 0. \end{cases} \tag{4.2}$$

In practice, the proportion $\frac{\sum I\left(w(h_{(ij),k}=1)\right)}{\binom{n}{2}\times m}$ is small. We aim to identify the latent patterns of $\{\beta_{(ij),k} \neq 0\}$ in B and G. In the human brain, structural-functional association patterns are often present in a multi-level dense structure, where FC-SI pairs with strong associations are primarily distributed in dense bipartite subgraphs in B. Further, the FCs in the dense bipartite subgraphs are inherently composed of dense clique subgraphs in the second-level graph G, while sparsely distributed in the rest of G. Formally speaking, in a *dense* bipartite subgraphs consisting of subsets of SIs and FCs, the possibility of strong FC-SI association within these subgraphs is much elevated than the rest FC-SI pairs outside of the subgraphs. This concept also applies to the dense clique subgraphs in G. We characterize this multi-level *dense* subgraph structure of FC-SI associations $\{\beta_{(ij),k} \neq 0\}$ as below:

$$\begin{aligned} &\text{Level 1: } B_c = (S_c, F_c; H_c) \text{ in } B, \\ &\qquad \text{where } \mathbb{P}\left(w(h_{(ij),k}) = 1 \mid h_{(ij),k} \in H_c\right) \gg \mathbb{P}\left(w(h_{(ij),k}) = 1 \mid h_{(ij),k} \notin H_c\right), \\ &\text{Level 2: } F_c \text{ unfolds into edges in } G_c = (V_c; F_c), \\ &\qquad \text{where } \mathbb{P}\left(w(f_{ij}) = 1 \mid i, j \in G_c\right) \gg \mathbb{P}\left(w(f_{ij}) = 1 \mid i, j \notin G_c\right). \end{aligned} \tag{4.3}$$

Level 1 indicates that FC-SI associations $\{\beta_{(ij),k} \neq 0\}$ are highly concentrated in a *dense* bipartite subgraph B_c, where SI measures in S_c are more likely to be associated with FC measures in F_c.

This may reflect SIs in a neural pathway specifically affect a group of FCs. Level 2 describes that F_c in B_c are not isolated edges in whole-brain connectome; instead, they often constitute an organized *dense* subgraph (G_c) as well. For simplicity, we refer to this multi-level subgraph as B_c. B_c characterizes that a subset of SI measures systematically impacts FC sub-networks. Neurologically, the SI-impacted sub-networks may be different from commonly known networks, such as the salience network or default mode network.

Remarks: If edges $\{h_{(ij),k}|w(h_{(ij),k}) = 1\}$ are randomly distributed in B, classic methods like FDR and sCCA can be applied to identify those edges with strong FC-SI association. However, if $\{h_{(ij),k}|w(h_{(ij),k}) = 1\}$ are not randomly distributed, the systematic FC-SI association patterns can be characterized as B_c. By constraining $\{\beta_{(ij),k} \neq 0\}$ in organized sub-networks B_c, we can critically improve the sensitivity of statistical inference while rigorously controlling false positive findings, which is to be shown in the following section.

4.2.3 B_c suppressing false positive findings

In this section, we demonstrate that compared to testing individual associations $\{\beta_{(ij),k}\}$, the identification of *dense* FC-SI association sub-network B_c can drastically reduce the false positive rate. Without loss of generality, we assume that false positive associations are randomly distributed. This assumption is commonly used in neuroimaging statistics and is plausible neurobiologically (Margulis and Sagan, 2000). Specifically, let B^* denote an FC-SI association bipartite (binary) graph that only includes false positive associations (i.e., no true positive association presents), and $\{\beta^*_{(ij),k} \neq 0|\beta_{(ij),k} = 0\}$ denote false positive edges. In B^*, the false positive rate is $p_1 = \frac{\sum_{ijk} I(\beta^*_{(ij),k} \neq 0)}{|S||F|}$. The subset of FC variables in $\{\beta^*_{(ij),k} \neq 0\}$ unfolds into false positive edges in G

with a false positive rate $p_2 = \frac{\sum_{ij \in G^*} I(\beta^*_{(ij),k} \neq 0)}{\binom{|V|}{2}}$. We show in Lemma 4.1 that the probability of $\{\beta^*_{(ij),k} \neq 0\}$ constituting a dense subgraph B_c is approximately zero.

Lemma 4.1. *Assume that B_c is observed from a random multi-level binary graph with a bipartite graph in level 1: $B^*(S, F; H)$ with connection density p_1 and a unipartite graph in level 2: $G^*(V; F)$ with connection density p_2. Also, suppose that B_c is a multi-level subgraph that has: (1) Edge density in B_c (level 1) with $\frac{\sum_{ijk \in B_c} w(h_{(ij),k})}{|S_c||F_c|} \geq \gamma_1 \in (p_1, 1)$; (2) Edge density in G_c (level 2) with $\frac{\sum_{ij \in G_c} w(f_{ij})}{\binom{|V_c|}{2}} \geq \gamma_2 \in (p_2, 1)$. Furthermore, let $m_0, n_0 = \Omega(\max\{m^\epsilon, n^\epsilon\})$ for some $0 < \epsilon < 1$, where Ω denotes a loose lower bound. Then for sufficiently large m, n with $\zeta(\gamma_1, p_1)m_0 \geq 4 \log n(n-1)$, $\zeta(\gamma_1, p_1)n_0(n_0-1) \geq 16 \log m$, and $\zeta(\gamma_2, p_2)n_0 \geq 4 \log n$, we have*

$$\begin{aligned} &\mathbb{P}\left(|S_c| \geq m_0, |F_c| \geq \binom{n_0}{2}, |V_c| \geq n_0\right) \\ &\leq 2mn^2(n-1) \cdot \exp\left(-\frac{1}{8}\zeta(\gamma_1, p_1)\, m_0 n_0 (n_0-1) - \frac{1}{4}\zeta(\gamma_2, p_2)\, n_0^2\right), \end{aligned} \tag{4.4}$$

where $\zeta(a, b) = \left\{\frac{1}{(a-b)^2} + \frac{1}{3(a-b)}\right\}^{-1}$.

Proof. The proof of Lemma 4.1 is provided in Appendix 4B.

Lemma 4.1 states that the probability that FC-SI false positive associations $\{\beta_{(ij),k} \neq 0\}$ compose into a dense multi-level subgraph B_c exponentially converges to zero with size $|S_c||F_c|$ and $|V_c|^2$. By De Morgan's law, an extracted multi-level subgraph B_c is unlikely to be false positive if B_c is both dense and reasonably-sized (e.g., 10 by 10 $|S| \times |F|$). This conclusion is also applicable to B^* including both false positive and true positive associations because Lemma 4.1 holds for any subgraph. Therefore, we are motivated to detect maximal dense multi-level

subgraphs because it naturally prohibit false positive findings while improving sensitivity (see Section 2.5).

4.2.4 Multi-level sub-network extraction

Inspired by Lemma 4.1, we propose an objective function to extract the dual dense multi-level sub-graphs $\{B_c\}$ from graphs B and G. To begin with, we first obtain the statistical inferential results for $\{\beta_{(ij),k}\}$, for example, p-values or the test statistics, which are then stored in an *inference matrix* $\boldsymbol{A} = \{a_{(ij),k}\}$. Each entry $a_{(ij),k}$ can be calculated based the inferential results (e.g., $-\log p_{(ij),k}$). In practice, p value-based inference is popular in high-dimensional data analysis, which is required by commonly used tools including multiple versions of False Discovery Rate (FDR) control, Manhattan plot, and combined p-value methods. Regardless, MOAT is a general approach that is applicable to any statistical inferential results stored in $\boldsymbol{A}$ as the input of the model.

For targeted multi-level subgraphs, we define a weighted matrix $\boldsymbol{U} \in \mathbb{R}^{\binom{n}{2}\times m}$ based on $\boldsymbol{A}$ and B_c (Chen et al., 2023):

$$\mathbf{U} = \boldsymbol{A} * \Delta, \quad \text{that is,} \quad u_{(ij),k} = a_{(ij),k} \cdot \delta_{(ij),k}, \tag{4.5}$$

where the symbol"$*$" denotes the Hadamard matrix product and $\delta_{(ij),k} = 1$, if $i, j, k \in B_c$; 0 otherwise. It is evident that $\mathbf{U}$ is contingent upon the multi-level network structure underlying graph B and G, as outlined in the previous section. Now, we proceed with the extraction of $\{B_c\}$ based on $u_{(ij),k}$ and the edge weight $w(f_{ij})$ in G by optimizing the core objective function

proposed in this article:

$$\underset{S_c \subseteq S, V_c \subseteq V}{\arg\max} \sum_c \frac{\sum_{ijk \in B_c} u_{(ij),k}}{\left(|S_c| \binom{|V_c|}{2}\right)^{\lambda_1/2}} + \frac{\sum_{ij \in G_c} w(f_{ij})}{|V_c|^{\lambda_2}}, \tag{4.6}$$

where λ_1, λ_2 are tuning parameters controlling the densities of B_c and G_c, respectively. When $\lambda_1 = 0$, the first term in (4.6) is simply $f_1 = \sum_{i,j,k \in B_c} u_{(ij),k}$, which quantifies the magnitude of significantly associated FC-SI pairs within sub-graph B_c. When $\lambda_1 = 2$, the first term becomes $f_2 = \frac{\sum_{i,j,k \in B_c} u_{(ij),k}}{|S_c|\binom{|V_c|}{2}}$, which is a popular definition for the connection density. Optimizing f_2 alone may result in the detection of an over-dense B_c with a very small size, such as a singleton, while optimizing f_1 alone can often lead to an oversized B_c that sacrifices the density. λ_2 plays a similar role for the sub-graph G_c. The two denominators in (4.6) refer to the number of variables in B_c and G_c, and thus are ℓ_0 penalties. Previous studies have demonstrated that ℓ_0 penalties are effective in detecting subgraphs with balanced size and density (Chen et al., 2023; Shabalin et al., 2009; Wu et al., 2021). Furthermore, λ_1 and λ_2 can be objectively selected by Kullback–Leibler (KL) divergence with the reference distribution based on the random graph. We provide detailed constructions of the reference distribution and selection procedures in Appendix 4A.

Directly optimizing the objective function (4.6) is a challenging task since it is a nondeterministic polynomial (NP) problem. To tackle this issue, we propose a computationally efficient greedy peeling algorithm that provides an approximate solution to the optimization problem. We build upon the greedy peeling algorithms used for sub-community detection in single-level bipartite graphs (Chekuri et al., 2022; Wu et al., 2021) and extend them to multi-level bipartite graphs to extract dense $\{B_c\}$. The greedy peeling algorithm is formally presented in Algorithm 5 below. Algorithm 5 can automatically account for the dependence between SI predictors themselves

Algorithm 5 Optimization of objective function (4.6)

Input: $B = (S, F; H)$, $G = (V, E)$, λ_1, λ_2
output: $\{B_c\}$

Define **function** [B_c density(B_c)]= greedy peeling_MOAT [B G λ_1 λ_2 $\{t_z\}_1^Z$]:
for $t = t_1, t_2, \ldots, t_Z$ **do**
  Assign $S_c^1 \leftarrow S, F_c^1 \leftarrow F$
  for $q = 1, 2, \ldots, m + \binom{n}{2} - 1$ **do**
    Let $\tau \in S_c^{(q)}$ be the node with smallest degree:
    $\tau = \arg\min_{\tau' \in S_c^{(q)}} \deg_X \left(\tau'; S_c^{(q)}, F_c^{(q)}\right)$,
    Let $\phi \in F_c^{(q)}$ be the node with smallest degree:
    $\phi = \arg\min_{\phi' \in F_c^{(q)}} \deg_Y \left(\phi'; S_c^{(q)}, F_c^{(q)}\right)$
    if $\sqrt{d} \deg_X \left(\tau; S_c^{(q)}, F_c^{(q)}\right) \leq \frac{1}{\sqrt{d}} \deg_Y \left(\phi; S_c^{(q)}, F_c^{(q)}\right)$ **then**
      $S_c^{(q+1)} \leftarrow S_c^{(q)} / \{\tau\}$ and $F_c^{(q+1)} \leftarrow F_c^{(q)}$
    else
      $S_c^{(q+1)} \leftarrow S_c^{(q)}$ and $F_c^{(q+1)} \leftarrow F_c^{(q)} / \{\phi\}$
    end if
    Construct $G = (V; F)$
    Assign $V_c^1 \leftarrow V$
    for $p = 1, 2, \ldots, n - 1$ **do**
      Let $a \in V_c^{(p)}$ be the node with smallest degree:
      $a = \arg\min_{a' \in V_c^{(p)}} \deg_Z(a'; V_c^{(p)})$
      $V_c^{(p+1)} \leftarrow V_c^{(p)} / \{a\}$
    end for
    Output $G_c[V_c^{(p)}; F_c^{(p)}]$ that maximizes $\frac{\sum_{i,j \in G_c} w(f_{ij})}{\left|V_c^{(p)}\right|^{\lambda_2}}$
    Update $F_c^{(q+1)} \leftarrow F_c^{(p)}$
  end for
  Output $B_c^{(t)}$ that maximizes (4.6) among $\{B_c^{(q)}\}_{q=1}^{m+\binom{n}{2}-1}$
end for
Output B_c that maximizes (4.6) among $\{B_c^{(t)}\}_{t \in (t_1, \ldots, t_Z)}$
end function

[B_c density(B_c)]= greedy peeling_MOAT [B G λ_1 λ_2 $\{t_z\}_1^Z$]
while density(B_c) $>$ density(B) **do**
  Fill median of $\{w(h_{(ij),k}), \forall i, j, k\}$ into $\{w(h_{(ij),k}) | i, j, k \in B_c\}$
  Obtain updated B^* and G^* with filled $\{w(h_{(ij),k})^*, \forall i, j, k\}$
  [B_c density(B_c)]= greedy peeling_MOAT [B^* G^* λ_1 λ_2 $\{t_z\}_1^Z$]
end while
Output $\{B_c\}$

because correlated SI predictors tend to be associated with highly overlapped subsets of FC outcomes. Following this rationale, MOAT subgraph extraction procedure also accounts for the dependence within the FC outcomes.

In Theorem 4.2.4, we establish the consistency of the multi-level subgraph detection. Specifically, the solution to optimize the objective function (4.6) provides a consistent estimate for true community structure of $\{B_c\}$ (the collection of edge-induced sub-networks), and the probability of incorrect edge edge assignment for $\{B_c\}$ converges to zero when the sample size $D \to \infty$.

1 (**Consistency of subgraph detection**). let $\mathcal{M}^*$ be the true membership of edges (the community index of edges falling in the dense sub-network B_c. Let $\hat{\mathcal{M}}$ be the edge membership for $\hat{B}_c$ estimated by optimizing function (4.6) using Algorithm 5. Then for an arbitrarily small ϵ and sample size $D \to \infty$, we have

$$\mathbb{P}(|\mathcal{M}^* - \hat{\mathcal{M}}| < \epsilon) = 1$$

Proof. Proof of Theorem 4.2.4 is provided in Appendix 4B.2.

The computational complexity of Algorithm 5 is $\mathcal{O}\Big(\mathcal{C}|V|(|S|+|F|)\Big)$, where $\mathcal{C}$ is a constant determined by the number of the grid searching, $|V| = n$ is the number regions, $|S| = m$ is the number of SI measures, and $|F| = \binom{n}{2}$ is the number of FC measures. Therefore, MOAT is computationally efficient. For example, the computational time is around 3 minutes for a structural-functional neuroimaging dataset with 5000 FCs and 750 SIs, using MATLAB 2021 with 6-Core Intel Core i5 of CPU and 32 GB of RAM.

4.2.5 Inference for $\hat{B}_c$

Recall that the goal of this study is to identify subsets of SIs and FCs that form systematic association patterns, i.e., the multi-level sub-networks $\{\hat{B}_c\}$. Given extracted $\{\hat{B}_c\}$ in section 4.2.4, we perform network-level inference to assess their statistical significance. This procedure includes both the global level and individual network level inference (Chen et al., 2023; Zhang et al., 2023).

Global-level inference. The global-level inference is to test the overall existence of multi-level sub-networks B_c:

$$\begin{aligned} &\mathbb{H}_0 : \text{No FC-SI systematic association (i.e., } B_c\text{) exists in } B; \\ &\mathbb{H}_1 : \text{At least one } B_c \text{ exists in } B. \end{aligned} \tag{4.7}$$

The significance of the deviation from $\mathbb{H}_0$ (e.g., p-value) can be directly calculated by the probability formula (4.4) derived in Lemma 4.1. The probability is the upper bound of the density and size of B, which can be conveniently used as the probability for the exact test. We reject $\mathbb{H}_0$ if this probability is a than the pre-specified significance level.

Individual-network-level inference. The individual-network-level inference aims to identify the multi-level sub-networks $\{\hat{B}_c\}$ with multiple corrections. Specifically, we assess the statistical significance of each B_c by testing:

$$\begin{aligned} &\mathbb{H}_0^c : B_c \text{ is a random multi-level subgraph;} \\ &\mathbb{H}_1^c : B_c \text{ is not a random multi-level subgraph, reflecting systematic FC-SI associations.} \end{aligned} \tag{4.8}$$

Since the asymptotic inference for B_c as an object is challenging, we adopt the commonly used cluster-wise inference (i.e., permutation tests) in neuroimaging analysis to test $\{\hat{B}_c\}$ with multiple correction. In the permutation tests, we consider $\hat{B}_c$ as a "cluster" in a multi-level graph space.

To improve inference accuracy, we propose a tailored test statistic $T(\hat{B}_c)$ based on Lemma 4.1:

$$\mathcal{T}(\hat{B}_c) = \exp\left(-\frac{1}{8}\zeta\left(\gamma_1, p_1\right) m_0 n_0 (n_0 - 1) - \frac{1}{4}\zeta\left(\gamma_2, p_2\right) n_0^2\right), \tag{4.9}$$

where $\zeta(a,b) = \left\{\frac{1}{(a-b)^2} + \frac{1}{3(a-b)}\right\}^{-1}$, and $\gamma_1, \gamma_2, p_1, p_2$ represent connection densities defined in Lemma 4.1. The test statistic $T(\hat{B}_c)$ reflects the combinatorial probability for B_c with a given size and density under $\mathbb{H}_0^c$. shuo says: need to elaborate more on $T(\hat{B}_c)$.

We formally present our proposed *network-based permutation test* for the significance of each extracted $\hat{B}_c$ in Algorithm 6, where $\|G\|_0$ stands for the number of edges in the graph G. The permutation procedure outlined in Algorithm 6 is effective in simulating the null distribution of the test statistic $T(\hat{B}_c)$. Therefore, FWER can be controlled effectively, yielding a corrected p-value for each extracted $\hat{B}_c$.

Algorithm 6 Assess the significance of $\{\hat{B}_c\}_{c=1,\ldots,C}$

Input: $\{\hat{h}_{(ij),k}\}, \hat{C} > 0, \{\hat{B}_c\}, \alpha$

output: FWER-controlled significance values for each $\hat{B}_c$

1. Choose a sound cut-off $\hat{r}$ and binarize graph $B[\hat{r}]$:$B[\hat{r}]_{(ij),k} = I(a_{(ij),k} > \hat{r})$
2. Estimate the distinct edge densities for B and G respectively:

$$\text{Overall density}: \hat{\mu}_1 = \frac{\sum_{1\le i\le j\le n, 1\le k\le m} I(a_{(ij),k} > \hat{r})}{m\binom{n}{2}}, \quad \hat{\mu}_2 = \frac{\sum_{1\le i\le j\le n} w(f_{ij})}{\binom{n}{2}};$$

$$\text{Within-subgraph density}: \hat{\gamma}_1 = \frac{\sum_{(i,j,k)\in \hat{B}_c} I(a_{(ij),k} > \hat{r})}{||\hat{B}_c||_0}, \quad \hat{\gamma}_2 = \frac{\sum_{(i,j)\in \hat{G}_c} w(f_{ij})}{||\hat{G}_c||_0}.$$

3. Calculate test statistic $\mathcal{T}_0(\hat{B}_c)$ for the current $\hat{B}_c$
4. Shuffle group labels of data P times and implement MOAT on each shuffled graph
5. Store the maximal test statistic for each simulation $l = 1, \ldots, L$:

$$\mathcal{T}_l = \sup_{c=1,\ldots,\hat{C}^l} \exp\left(-\frac{1}{8}\zeta\left(\gamma_1^l, p_1^l\right)\left\|\hat{B}_c^l\right\|_0 - \frac{1}{4}\zeta\left(\gamma_2^l, p_2^l\right)\left(\left\|\hat{G}_c^l\right\|_0\right)^2\right)$$

6. Calculate the percentile of $\mathcal{T}_0(\hat{B}_c)$ in $\{\mathcal{T}_l\}_{l\in(1,\ldots,L)}$ as the FWER q-value and reject $\mathbb{H}_0$ if $q < \alpha$, where α is a pre-specified significant level

Evaluating the joint effect of multiple SIs on FCs. The resulting significant sub-networks $\{\hat{B}_c\}_{c\in(1,\ldots,C)}$ by MOAT are compatible with a variety of existing multivariate analysis tools. For example, given estimated $\{\hat{B}_c\}$, we can further examine the joint effects of multiple SI variables on FC sub-networks as performing the classic multiple regression analysis. Specifically, we can apply CCA on each selected $\hat{B}_c$ by building a set of linear combinations named $\mathcal{U}$ and $\mathcal{V}$, corresponding to the set of SI variables $\boldsymbol{X} = \{x_k\}_{k=1}^{|S_c|}$ and the set of FC variables $\boldsymbol{Y} = \{y_i\}_{i=1}^{|F_c|}$, respectively. Assuming $k < i$, we construct $\mathcal{U}_k = a_{k1}\boldsymbol{X}_1 + a_{k2}\boldsymbol{X}_2 + \ldots + a_{k|S_c|}\boldsymbol{X}_{|S_c|}$ for $k \in (1, \ldots, |S_c|)$ and $\mathcal{V}_i = b_{i1}\boldsymbol{Y}_1 + b_{i2}\boldsymbol{Y}_2 + \ldots + b_{i|F_c|}\boldsymbol{Y}_{|F_c|}$ for $i \in (1, \ldots, |S_c|)$. Defining $(\mathcal{U}_j, \mathcal{V}_j)$ as the j-th canonical variate pair, we can then search for $\mathcal{U}_j$ and $\mathcal{V}_j$ that maximize the canonical correlation $\rho_j^* = \frac{\mathrm{cov}(U_j, V_j)}{\sqrt{\mathrm{var}(U_j)\,\mathrm{var}(V_j)}}$. Last, we scrutinize the joint effects of SIs on FC sub-networks within the selected $\hat{B}_c$ based on the top largest $\{\rho_j^*\}_{j<|S_c|}$.

Additionally, once $\{\hat{B}_c\}$ are detected, one may further conduct low-rank techniques on each

$\hat{B}_c$ to estimate the final effect sizes, for example, low-rank regression models for high-dimensional feature selections (Kong et al., 2019; Vounou et al., 2010; Wang et al., 2012; Zhu et al., 2014). Alternatively, we can also apply advanced regression models, such as tensor regression models (Zhou et al., 2013) and sparse multi-task regression for brain imaging data Wang et al. (2011). In short, further applying these low-rank models on selected $\{\hat{B}_c\}$ allows us to validate the rigorous statistical association of FC-SI pairs selected by MOAT, where we have already ensured the *systematic* association patterns between these selected pairs as well as the topological structures of brain sub-networks composed by the selected FCs.

4.3 Simulation

In the simulation study, we probed whether MOAT can extract informative sub-graphs $\{\hat{B}_c\}$ from the multi-level graph B and G with high accuracy and multiple corrections. We evaluate MOAT to finite-sample simulation data under various conditions (e.g., different sample sizes and effect sizes) with comparisons to various commonly used biclustering methods and sCCA-based methods.

4.3.1 Synthetic data

For the synthetic data, we aim to construct a matrix $\boldsymbol{A}_{500\times4950}$ that stores the partial correlation values between 500 SIs and 4950 FCs. The FC measures are calculated based on a brain network with 100 regions, resulting in $\binom{100}{2} = 4950$ pairwise connectivity values. $\boldsymbol{A}_{500\times4950}$ governs the weight of edge variable H in the bipartite graph $B = (S, F; H)$ by $w(h_{(ij),k}) = I(a_{(ij),k} > r)$, where r is a threshold for correlation strength. Two sub-networks

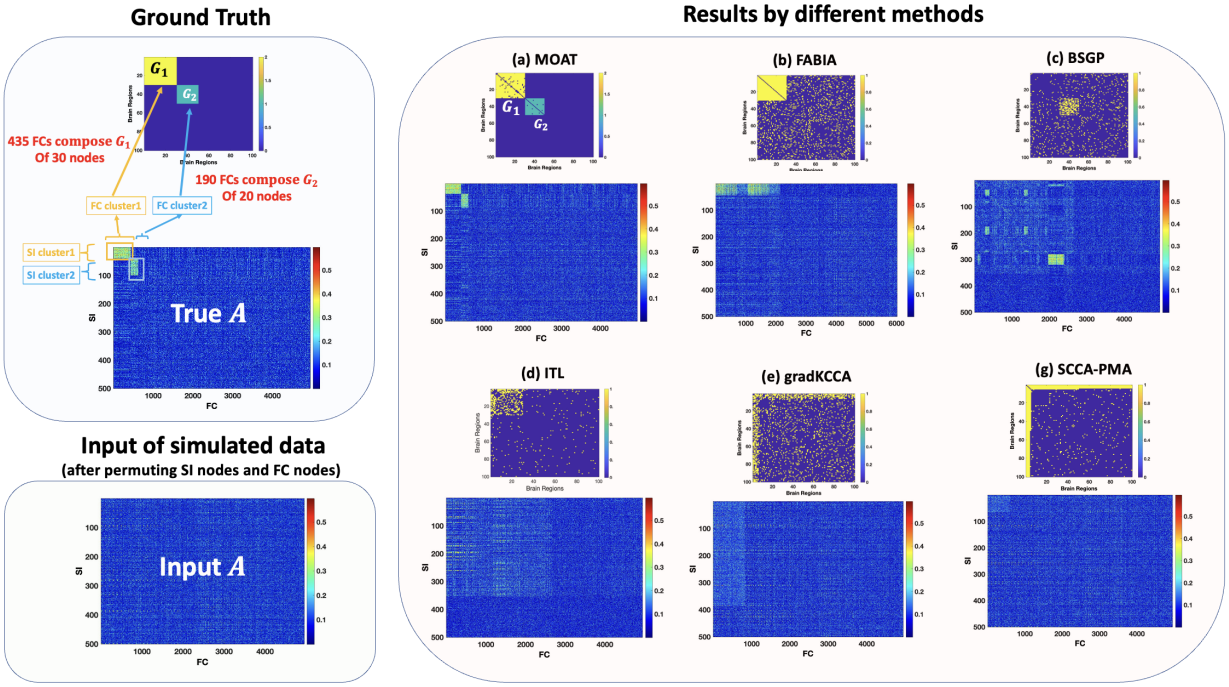


Figure 4.3: *Application of MOAT and comparative methods on a synthetic partial correlation matrix* $\boldsymbol{A}_{500\times4950}$*, which represents the associations between* 500 *SI measures and* 4950 *FC measures (obtained from pair-wise connections between 100 brain regions). Two sub-networks* B_1 *and* B_2 *with high FC-SI partial correlations are generated within* B *and* G*. We applied MOAT and comparative methods to identify* B_1 *and* B_2*. The result panel on the right shows that MOAT and biclustering-based methods are more accurate in recovering sub-network patterns than sCCA-based methods since they incorporate cluster/network information. When jointly evaluating true positive rate (TPR) and true negative rate (TNR), MOAT outperforms all other methods since TPR and TNR rely heavily on accurate sub-network extraction and inference, and MOAT is specifically designed to capture and characterize dense sub-network structures by leveraging graph combinatorial theories.*

B_1 and B_2 with high FC-SI partial correlations ρ_1, ρ_2 are generated within B and G (see graph illustration in Figure 4.3 for a visual representation of these two sub-networks.). Specifically, B_1 consists of 40 SI measures and 435 FC measures that compose G_1 of 30 brain regions; B_2 consists of 60 SI measures and 190 FC measures that compose G_2 of 20 brain regions. Next, we vary $\rho_1, \rho_2 > \rho_0$ to emulate different effect sizes, where $\rho_0 = 0.2$ is set to be the partial correlation of the remaining FC-SI edges out of informative sub-networks B_1, B_2. D is previously used to denote the sample size of the data. Three configurations of $(\rho_1, \rho_2; D)$ are simulated: $(0.55, 0.60; 200)$, $(0.60, 0.45; 300)$, and $(0.70, 0.40; 400)$. We simulate 500 repeated data sets $\{\boldsymbol{A}^l\}_{l\in(1,\ldots,500)}$ for each configurations.

For each simulated $\boldsymbol{A}^l$, we apply MOAT to estimate the FC-SI associated multi-level sub-networks $\hat{B}_c$ and perform the proposed network-based permutation test. We benchmark MOAT against several popular (i) biclustering methods that are commonly used for community detection (ii) sCCA that can identify and measure the associations between two canonical/latent types of variables. Specifically, for biclustering methods, we adopt Bipartite Spectral Graph Partitioning (BSGP) (Wieling and Nerbonne, 2009), Information Theoretic Learning (ITL) (Erdogmus, 2002), and Factor Analysis for Bicluster Information Acquisition (FABIA) (Hochreiter et al., 2010); for sCCA, we adopt gradKCCA, a Large-Scale Sparse Kernel Canonical Correlation method proposed by Uurtio et al. (2019), and sCCA through a penalized matrix decomposition (sCCA-PMA) proposed by Witten et al. (2009).

$(\rho_1, \rho_2; \mathrm{n})$	measures	metrics	MOAT	BGSP	ITL	FABIA	gradKCCA	SCCA-PMA
(0.55, 0.60; 200)	SI	TPR	1 (0)	0.4767 (0.1357)	0.4010 (0.4169)	0.5768 (0.1019)	0.3450 (0.0376)	0.4309 (0.0147)
		TNR	0.9844 (0.0129)	0.9492 (0.0290)	0.9560 (0.0430)	0.7243 (0.0506)	0.4793 (0.0175)	0.6119 (0.0152)
	FC	TPR	0.9358 (0.0456)	0.4549 (0.0512)	0.4547 (0.0498)	0.8291 (0.0718)	0.4024 (0.0065)	0.5109 (0.0032)
		TNR	1 (0)	0.9462 (0.0129)	0.8683 (0.1423)	0.8288 (0.0263)	0.5547 (0.0076)	0.6643 (0.0047)
	SI-FC pair	TPR	0.9358 (0.0456)	0.4090 (0.0184)	0.3891 (0.0267)	0.4275 (0.0578)	0.3417 (0.0138)	0.4110 (0.0038)
		TNR	0.9844 (0.0129)	0.9082 (0.0333)	0.7352 (0.3918)	0.6491 (0.0814)	0.4528 (0.0911)	0.5834 (0.0124)
(0.60, 0.40; 300)	SI	TPR	1 (0)	0.4760 (0.1393)	0.3734 (0.3464)	0.5165 (0.1037)	0.3790 (0.0336)	0.3202 (0.0138)
		TNR	0.9761 (0.0253)	0.9540 (0.0309)	0.9547 (0.0786)	0.6953 (0.0636)	0.4350 (0.0168)	0.5375 (0.0142)
	FC	TPR	0.8815 (0.1480)	0.4571 (0.0572)	0.3557 (0.0542)	0.8048 (0.0097)	0.3888 (0.0069)	0.5103 (0.0031)
		TNR	1 (0)	0.9480 (0.0126)	0.9026 (0.1149)	0.8291 (0.0035)	0.4734 (0.0072)	0.5646 (0.0073)
	SI-FC pair	TPR	0.8815 (0.1480)	0.4101 (0.0245)	0.3552 (0.0262)	0.4069 (0.0127)	0.3492 (0.0168)	0.4047 (0.0035)
		TNR	0.9761 (0.0253)	0.9045 (0.0352)	0.7797 (0.3163)	0.6694 (0.0083)	0.3361 (0.0112)	0.5907 (0.0104)
(0.70, 0.45; 400)	SI	TPR	1 (0)	0.4649 (0.1278)	0.3507 (0.3737)	0.5677 (0.0094)	0.3956 (0.0512)	0.4009 (0.0188)
		TNR	0.9800 (0.0202)	0.9443 (0.0549)	0.9493 (0.1036)	0.7533 (0.0742)	0.4423 (0.0231)	0.6255 (0.0186)
	FC	TPR	0.9073 (0.1098)	0.4621 (0.0731)	0.3831 (0.0938)	0.8094 (0.0502)	0.3940 (0.0181)	0.5803 (0.0031)
		TNR	1 (0)	0.9112 (0.0047)	0.8568 (0.1952)	0.8073 (0.0327)	0.4245 (0.0175)	0.6479 (0.0056)
	SI-FC pair	TPR	0.9073 (0.0202)	0.4248 (0.0615)	0.3378 (0.0572)	0.8048 (0.0097)	0.4212 (0.0246)	0.4917 (0.0042)
		TNR	0.9800 (0.0202)	0.9055 (0.0442)	0.7939 (0.4354)	0.6730 (0.0996)	0.3317 (0.0328)	0.5660 (0.0238)
Computational Time (MATLAB, 6-Core Intel Core i5, 32 GB RAM)			2min	5min	22min	2min	1min	2min

Table 2: *Inference results of MOAT and comparative methods under different settings.* ρ_1, ρ_2 *denote the partial correlations between SI and FC measures in the dense sub-networks* B_1 *and* B_2*, respectively, and* D *denotes the sample size. We summarize the means (standard deviations) of the TPR and TNR for identified SIs, FCs, and FC-SI pairs based on 500 repeated simulations. The results show that MOAT outperforms the compared methods under different scenarios, indicating its effectiveness in identifying the correct dense sub-networks with true positive FC-SI associated pairs.*

4.3.2 Performance evaluation

We evaluate method performance by assessing the deviation of the estimated $\hat{B}_c$ from true B_c at both node-level and edge-level from three perspectives: SI-wise, FC-wise, and FC-SI pair-wise. The $\hat{B}_c$ versus true B_c differences from each perspective are evaluated using true positive rate (TPR) and true negative rate (TNR). Table 2 summarizes the performance of different methods under all settings. We also provide a 2D visualization of the detected sub-networks by each method on a simulated $\boldsymbol{A}_{500\times4950}$ in Figure 4.3, using $(\rho_0, \rho_1, \rho_2) = (0.20, 0.40, 0.35)$ and sample size $D = 200$. Under this setting, we observe that MOAT performs the best in terms of controlling TPR, TNR, and network pattern recovery, followed by FABIA, BGSP, ITL, gradKCCA, and sCCA-PMA.

Specifically, according to Table 2 , MOAT and biclustering-based methods can recover sub-network patterns more accurately than sCCA-based methods as the cluster/network information is utilized. When the partial correlations are set relatively larger (and so are the effect sizes), MOAT tends to recover denser sub-networks and include fewer false-positive FC-SI edges in the estimated sub-networks, which leads to the extraction closer to the true network. The biclustering-based methods also perform better with increased effect sizes. In contrast, under different effect sizes, sCCA-based methods yield more similar results but fail to recognize the network structure as they incorporate no graph information into the model. This results in less accurate results when jointly evaluating TPR and TNR. The biclustering-based methods, on the other hand, yield very high TNR, but at the expense of low TPR. Overall, MOAT is most robust to both false-positive and false-negative noise under different settings when jointly assessing TPR and TNR. This is because MOAT allows tuning between sub-network size and density in the objective function, which is

equivalent to imposing an l_0 penalty term. Moreover, the TPR and TNR heavily rely on accurate sub-network extraction and inference, which our new statistical inference methods can better capture and characterize covariate-related sub-networks. In summary, MOAT outperforms other biclustering and sCCA methods with well-controlled TPR and TNR thanks to its tailed community detection modeling and inference.

4.4 Study of FC-SI associations in brain connectome data

4.4.1 UK Biobank sample and neuroimaging data

We aim to investigate the systematic effects of structural brain imaging measures on the functional connectome using the UK Biobank data (Sudlow et al., 2015). The UK Biobank is a vast biomedical database encompassing approximately half a million participants from the UK. Specifically, we utilize the following three sets of measurements from eligible individuals:

(i) SI measures: We collected 105 SI variables from 39 white matter integrity measures obtained from DTI data, which reflect the overall health and coherence of brain white matter. Additionally, we collected 66 cortical thickness measures that gauge the width of the gray matter of the human cortex. The white matter integrity is assessed by FAs from the DTI data, following a preprocessing procedure based on the ENIGMA DTI protocols (Jahanshad et al., 2013) and labeled based on the JHU ICBM DTI-81 Atlas (Mori et al., 2008; Smith et al., 2006). A complete list of the 39 regional white matter tracts can be found in Appendix 4C.3. The cortical thickness measures were obtained from T1 MRI and labeled based on the FreeSurfer atlas (Tustison et al., 2014).

(ii) FC measures: The functional connectome data were obtained from rs-fMRI data. We first performed rs-fMRI preprocessing for all participants and extracted the averaged time series of blood-oxygen-level-dependent (BOLD) signals for 246 functional brain regions based on the Brainnetome Atlas (Fan et al., 2016).(see details of imaging acquisition and fMRI preprocessing in Appendix 4C.1.). The whole-brain functional connectome was calculated by correlating BOLD signals between all regions, resulting in $\binom{246}{2} = 30,135$ region-pair FC variables.

(iii) Confounding variables: we adjusted four confounding variables including age, sex, educational level, and body mass index (BMI), which are commonly used in the existing neuroimaging literature (Agustí et al., 2018; Alfaro-Almagro et al., 2021; Bischof and Park, 2015; Miller et al., 2016).

Out of half a million participants in the UK Biobank, a total of 40,923 healthy individuals were found to have usable resting-state fMRI (rs-fMRI) data that passed quality control (Alfaro-Almagro et al., 2018). Among them, a subgroup of 4,242 individuals possessed complete data on the 105 SI measures. We focus on this cohort of 4,242 participants for our primary investigation, who have an average age of 61.46 ± 7.40 (years), sex (M/F: 2003/2239), educational level (years: 17.37 ± 3.92), and BMI (kg/m^2 : 26.35 ± 4.30).

4.4.2 Results

We applied MOAT to the multimodal imaging and demographic data for these qualified 4,242 UK Biobank participants. First, we obtained the FC-SI association inference matrix $\boldsymbol{A}_{105\times 30135}$ by 4.1. Each entry in $\boldsymbol{A}$ is $a_{(ij),k} = -\log(p_{(ij),k})$, where $p_{(ij),k}$ represents the p-value testing the

association between the k-th SI measure and the FC outcome between two brain regions i and j. Next, we performed the hard-thresholding sparsity constraint by setting $a_{(ij),k} = a_{(ij),k} I(a_{(ij),k} < \epsilon)$ for some positive integer ϵ (Zhang et al., 2023). Lastly, we applied the proposed greedy peeling algorithm 5 to the inference matrix $\boldsymbol{A}$, with tuning parameters $\lambda_1 = 1.25, \lambda_2 = 1.5$ which were selected by the KL divergence with a mixed Bernoulli distribution based on random B and G. Algorithm 5 identified one multi-level sub-network $\hat{B}_1 \in B$. Lastly, we performed the network-level statistical inference on $\hat{B}_1$ using Algorithm 6. The testing results showed that the systematic association pattern of $\hat{B}_1$ is statistically significant ($p < 0.0001$).

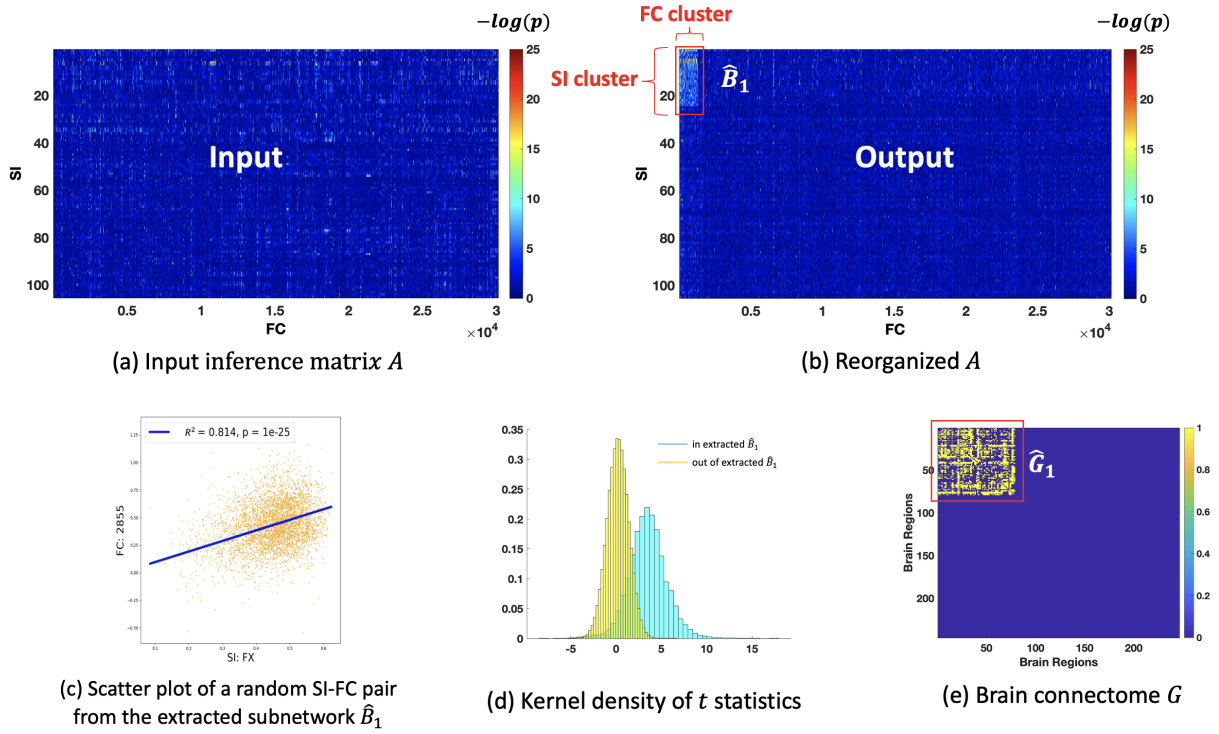


Figure 4.5: *Application of MOAT on a real neuroimaging dataset obtained from the UK Biobank. (a) shows a heatmap of the inference matrix* $\boldsymbol{A}_{105\times30135}$*. Each element is a* $-log(p)$ *value testing the association between SI and FC measures. (b) shows a heatmap of a reorganized* $\boldsymbol{A}$ *with element in the detected subgraph* $\hat{B}_1$ *pushed to the top left corner. Within the second level* G *composed by the selected FCs, there exists an organized subgraph* $G_1 \subset G$ *revealed in (e). As shown in (c) and (d), the FC-SI pairs within the identified* $\hat{B}_1$ *have significantly stronger associations compared to those outside of the network.*

Specifically, $\hat{B}_1$ comprised $|S_1| = 23$ SI measures and $|F_1| = 1316$ FC outcomes, as

highlighted in Figure 4.5(b). Furthermore, the extracted $|F_1|$ unfolded into a dense clique $\hat{G}_1 \in G$ consisting of $|V_1| = 79$ regions within the second-level brain functional connectome G, as illustrated in Figure 4.5(e). The FC-SI pairs within the identified $\hat{B}_1$ demonstrate significantly stronger associations compared to those outside of the network, as evidenced by the results (R^2 and t-statistics) shown in Figure 4.5 (c-d). Among the 23 extracted SI measures, three of them are the region-wise cortical thickness measures and the remaining 20 measures are FAs. The three cortical thickness measures capture the mean thickness of the parahippocampal, superior temporal, and cuneus gyrus. Among the extracted FA measures, the top four with the strongest FC associations are CST-R (corticospinal tract, right hemisphere), CST-L (corticospinal tract, left hemisphere), ICP (inferior cerebellar peduncles), and FX (fornix). More detailed information about the remaining 16 FA measures can be found in Appendix 4C.4. Figure 4.6 (left panel) illustrates the names of all 20 selected FAs. On the other hand, the 79 brain regions within the extracted FC sub-network are predominantly located in six cortices: frontal, subcortical, temporal, parietal, insular, and limbic. Their spatial distributions are visually demonstrated in Figure 4.6 (right panel). The FC subnetwork in $\hat{B}_1$ generally consists of several well-defined brain functional networks including temporo-frontal, somatomotor, ventral attention, frontoparietal, and (partial) default mode network (DMN). Overall, Figure 4.6 provides a 3D demonstration showcasing the systematic association patterns between the subsets of SIs and FCs revealed by MOAT. Notably, both the FC-SI sub-network ($\hat{B}_1$) and the brain functional sub-connectome ($\hat{G}_1$) exhibit well-organized topological structures.

We further applied CCA on the extracted sub-network $\hat{B}_1$ to quantitatively measure the canonical associations among the FC-SI pairs within $\hat{B}_1$. Results showed that the sample canonical correlations of the first three canonical variate pair in $\hat{B}_1$ were 0.81, 0.69, and 0.68 respectively.

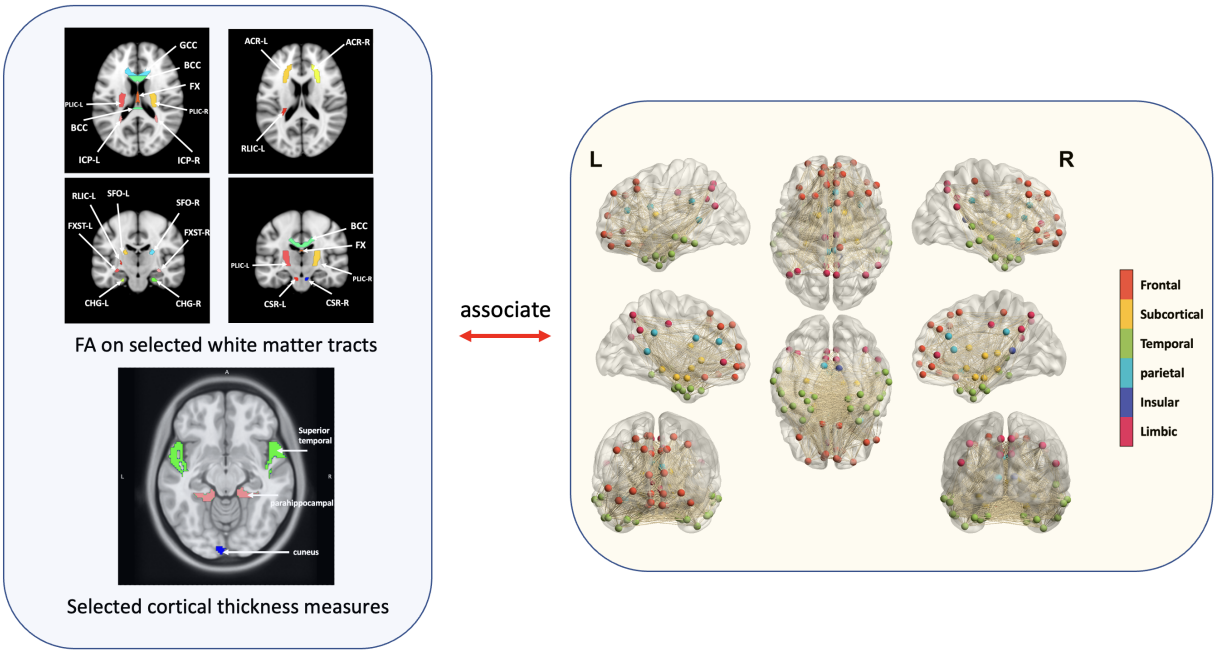


Figure 4.6: *Extracted FC-SI associated sub-networks by MOAT, adjusted for age, sex, educational level, and BMI. Specifically, the left panel shows the SI cluster in the extracted sub-networks, which contains 20 FA measures and 3 cortical thickness measures. The top four FA measures with remarkably stronger associations with the detected FC sub-network are CST-R, CST-L, ICP, and FX. The three selected cortical thickness measures correspond to the mean thickness of the parahippocampal, superior temporal, and cuneus gyrus. The right panel shows the detected FC sub-network that is strongly impacted by the 23 selected SIs. The detected FC sub-network covers 79 brain regions and composes several high-level cognitive brain networks including the default mode network, temporo-frontal network, somatomotor, ventral attention, and frontoparietal network.*

Next, we performed sparse CCA proposed by Witten et al. (2009) on the full graph B and G, given the ultra high dimensionality of data. The yielded sample canonical correlations of the first three Canonical variate pair were 0.14, 0.18, and 0.20, respectively. We observe that MOAT extracted a parsimonious set of FC-SI pairs with strong canonical correlations while simultaneously ensuring the *systematic* patterns of both the FC-SI associations and the brain connectomic network composed by the selected FCs.

In summary, the application of MOAT helps to unfold the complex yet systematical and strong interplay between subsets of structural and functional measures of the human brain. Our findings suggest i) FC-SI associations are highly concentrated in a subset of SIs and FC sub-networks rather than exhibiting a whole-brain diffuse distribution pattern; ii) several FC sub-

Table. 20 selected white matter tracts strongly associated with identified FC sub-network

Region	Abbreviation	Category	Function
Inferior cerebellar peduncle	ICP	Brainstem	coordination of movement and proprioceptive
Genu of corpus callosum	GCC	Commissural	information transmission; movement control; cognitive functions; vision
Body of corpus callosum	BCC	Commissural	
Fornix	FX	Association	memory formation and retrieval
Corticospinal tract R	CST-L	Brainstem	Motor signal transmission; voluntary muscle control
Corticospinal tract L	CST-R	Brainstem	
Posterior limb of internal capsule R	PLIC-R	Projection	sense; motor; information flow past the basal ganglia
Posterior limb of internal capsule L	PLIC-L	Projection	
Retrolenticular part of internal capsule L	RLIC-L	Projection	
Anterior corona radiata R	ACR-R	Projection	voluntary emotional expression; sensory input
Anterior corona radiata L	ACR-L	Projection	
Cingulum hippocampus R	CHG-R	Association	Affect; response selection in skeletomotor control; visuospatial processing and memory access
Cingulum hippocampus L	CHG-L	Association	
Fornix (cres)/stria terminalis R	FXST-R	Association	Memory; emotion, regulation of stress responses within the brain
Fornix (cres)/stria terminalis L	FXST-L	Association	
Superior fronto-occipital fasciculus R	SFO-R	Association	Auditory verbal and declarative
Superior fronto-occipital fasciculus L	SFO-L	Association	
Uncinate fasciculus R	UN-R	Association	Memory; language; emotion
Tapetum R	TAP-R	Commissural	information transmission; visual capabilities
Tapetum L	TAP-L	Commissural	

Table 3: *20 selected white matter tracts strongly associated with identified FC sub-network*

networks are primarily influenced by white matter integrity measures (refer to Table 3 for possible mapping relationships); iii) multiple SI measures jointly affect the overall FC outcomes based on MOAT-guided CCA analysis. While, on a high level, our results align well with previous medical findings (Chaddock-Heyman et al., 2013; Cheung et al., 2008; Corrigan et al., 2015; Pradat and Dib, 2009), MOAT reveals more refined patterns with improved spatial specificity and biological interpretability.

4.5 Discussion

Our newly developed approach, MOAT, offers a novel strategy to investigate the complex association patterns between multimodal neuroimaging data with matrix-outcomes (FC) and a vector of imaging predictors. MOAT deciphers the complex FC-SI association patterns in a multi-level graph model revealing the joint effect of a small set of SI predictors on FC sub-networks. The multi-level graph model can effectively reduce the number of parameters while preserving the spatial specificity of FCs and SIs. MOAT delivers findings in organized multi-level sub-networks largely suppressing individual false positive FC-SI associations (see Lemma 4.1 in section 4.2.3). We developed computationally efficient algorithms to extract multi-level sub-networks. We further showed the consistency of the MOAT method. In addition, we develop a tailored network-level inference approach to test extracted multi-level sub-networks while controlling multiplicity. Last, MOAT is also compatible with existing multivariate-to-multivariate analysis tools (e.g., CCA).

In our case study, we investigated the FC-SI associations based on a large sample and revealed systematic association patterns with neurological explanations. This may enhance our understanding of how the brain structure and function interactively work during resting states

and may lead to insights that can guide future cognitive and psychiatric therapy. Since UK biobank participants mainly consist of elder Caucasians, our conclusion may be limited. Further investigation and integrated analysis is required to gain more comprehensive understanding of the FC-SI associations. The software package for MOAT is available at https://github.com/TongLu-bit/multi-levelNetworks-MOAT.

Appendix : SCCN

2A. Spatial-contiguity constraints

2A.1. Formal definition of spatial-contiguity

In the main manuscript, we define the spatial contiguity between a pair of voxels within a region, say A, by an "infrastructure graph" $\mathcal{S}_A$. Specifically, $S_{ii'} = 1$ if $d_{ij} \leq \varepsilon$, and $S_{ii'} = 0$ otherwise, where $S_{ii'}$ is an entry in $\mathcal{S}_A$ and $d_{ii'}$ is an Euclidean distance between voxels i and i'. Built on this voxel-pair-level spatial contiguity, we can define the spatial contiguity for a set of voxels $U_0 \subset U$ within a region. U_0 is spatially contiguous if, for any two voxels $i, i' \in U_0$, there exists a path $L_{i,i'} \subset U_0$ connecting i and i', where a pair of voxels are connected in $L_{i,i'}$ if and only if they are spatially contiguous.

To define the spatial contiguity more rigorously, we can use a 3-tuple coordinate to indicate the spatial position of each node i as $(\mathfrak{x}_i, \mathfrak{y}_i, \mathfrak{z}_i)$. Based on this, we define an "infrastructure graph" $\mathcal{S}_A$ between all nodes in ROIs A. In $\mathcal{S}_A$, each node i is connected to spatially neighboring nodes $\{(\mathfrak{x}_i + \mathfrak{a}, \mathfrak{y}_i + \mathfrak{b}, \mathfrak{z}_i + \mathfrak{c}) : (\mathfrak{a}, \mathfrak{b}, \mathfrak{c}) \in \{-1, 0, 1\}^3 \backslash \{(0, 0, 0)\}$, and $(\mathfrak{x}_i + \mathfrak{a}, \mathfrak{y}_i + \mathfrak{b}, \mathfrak{z}_i + \mathfrak{c})$ belongs to ROI A. A subset U_0 of ROI A is referred to as being *spatially contiguous* if U_0 induces a connected subgraph in $\mathcal{S}_A$. For example, the subgraph $U_0 := \{V_1, \ldots, V_{14}\}$ in Figure A.1 is spatially contiguous, whereas $U_0' := \{V_{15}, \ldots, V_{21}\}$ is not. We define and interpret $\mathcal{S}_B$ for all

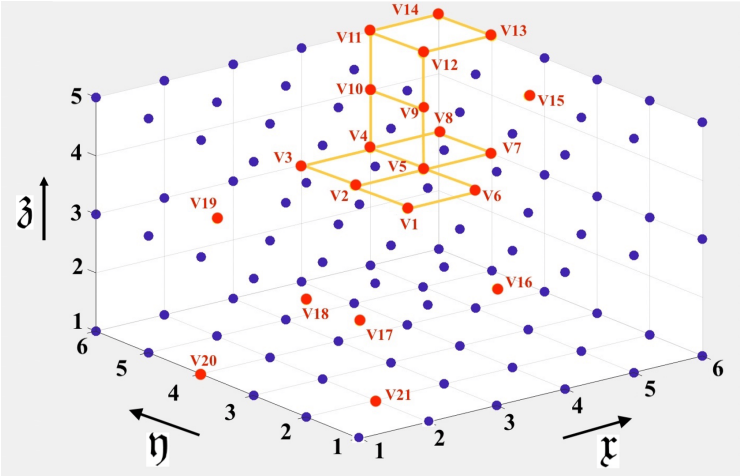


Figure A.1. *An illustration of the concept* spatial contiguity. *Based on the definition, the subgraph* $U_0 := \{V_1, \ldots, V_{14}\}$ *is spatially contiguous in this space, whereas* $U_0' := \{V_{15}, \ldots, V_{21}\}$ *is not.*

nodes in ROI B similarly.

2A.2. Implementation of spatial-contiguity constraints

To practically incorporate the within-region infrastructure graphs $\mathcal{S}_A$ and $\mathcal{S}_B$ into the between-region inference matrix $\mathbf{W}$, we adapt our model from a biclustering algorithm based on the use of *Singular Value Decomposition* (SVD) proposed by Cano et al. (2007). This biclustering algorithm simultaneously co-clusters two types of input variables based on the similarity of their joint interactions ($\mathbf{W}$ in our case), where a subset of variable 1 exhibits similar behavior across a subset of variable 2, and vice versa (Hartigan, 1972). Hence, applying SVD on $\mathbf{W} = \mathbb{U}\mathbf{\Sigma}\mathbb{V}^\top$ will generate targeted bi-clusters, i.e., sub-area pairs. Specifically, clustering $\mathbb{U}$ and $\mathbb{V}$ gives the partitions of region A and B, respectively. Furthermore, $\mathbb{V}$ is the eigenvectors of $\mathbf{W}^\top\mathbf{W}$, since $\mathbf{W}^\top\mathbf{W} = \mathbb{V}\mathbf{\Sigma}^\top\mathbf{\Sigma}\mathbb{V}^\top$, $\mathbf{\Sigma}^\top\mathbf{\Sigma}$ is a diagonal matrix, and $\mathbb{V}$ is an orthogonal matrix. Therefore, performing the eigen-decomposition on $\mathbf{W}^\top\mathbf{W}$ and clustering its yielded eigenvector will give the exact partitions of region B. Similarly, performing same procedures on $\mathbf{W}\mathbf{W}^\top$ will provide

partitions of region A. Now, we shall construct matrices that can incorporate the spatial-contiguity infrastructure graphs $\mathcal{S}_A$ and $\mathcal{S}_B$ by taking the element-wise product : $\mathbf{W_A} = \mathbf{W}\mathbf{W}^\top \odot \mathcal{S}_A$ and $\mathbf{W_B} = \mathbf{W}^\top\mathbf{W} \odot \mathcal{S}_B$.

2B. Within-region vFC association analysis

2B.1 Background

Data structure

We focus on testing the association between voxel-level functional connectivity (vFC) within a region and a clinical variable or experimental conditions. The within-region vFC for a participant can be represented by a binary or weighted adjacency matrix $\mathbf{Y_{n\times n}} = \{Y_{ii'}\}$, where each off-diagonal entry represents the connection between a pair of voxels within a region of interest (ROI) from a brain Atlas (e.g., Brainnetome Atlas by Fan et al. (2016)). Regression models are used to examine the association between the within-region vFC matrix $\mathbf{Y_{n\times n}}$ and a primary regressor of interest, denoted as x_1 (such as clinical status), while adjusting for other predictors. Specifically, the regression model can be expressed as:

$$\mathbb{E}(Y^s_{ii'}|\mathbf{X}^s) = \alpha_0 + x^s_1\beta_{ii'} + \mathbf{X}^s_{1\times(p-1)}\boldsymbol{\alpha},$$

where $\beta_{ii'}$ is the coefficient associated with x^s_1, and $\boldsymbol{\alpha}$ is a coefficient vector for the remaining covariates $\mathbf{X}^s_{1\times(p-1)}$ (e.g., age, sex, etc.) for subject s. Similar to the input matrix described in the paper, we summarize the significance levels of $\beta_{ii'}$ using a *within-region connectivity inference matrix* denoted as $\tilde{\mathbf{W}}_{n\times n}$. Also, similar to the definition in Section 2 of the main text, each entry in

$\tilde{\mathbf{W}}_{n\times n}$ can be computed as $\tilde{\mathbf{W}}_{ii'} = -\log p_{ii'}$, where $p_{ii'}$ is the p-value for $\beta_{ii'}$. The same definition applies to $\tilde{\mathbf{W}}_{m\times m}$ for an ROI with m regions.

Graph representation

To decipher the within-region voxel-pair connectome, we consider a unipartite graph $H = \{R, E\}$ that corresponds to the inference matrix $\tilde{\mathbf{W}}_{n\times n}$. The node set R represents the voxels within the selected ROI, where the size of the set is denoted as $|R| = n$. The edge set $E = e_{ii'}$ signifies the functional connections between voxels within the ROI, and its size is given by $|E| = n(n-1)/2$. We make the same assumption that, following spatial normalization and registration of the fMRI data, all subjects have a shared node set denoted as $R^s \equiv R$ for all $s \in [S]$.

Spatial contiguity

The definition of spatial contiguity for within-region voxels remains the same as described in the main text. Once again, we will utilize the "infrastructure graph" $\mathcal{S}A$ to represent the connections between all nodes within a chosen region, say ROI A, in order to incorporate spatial contiguity. Each entry $S_{ii'}$ in $\mathcal{S}_A$ serves as a spatial-adjacency indicator variable between voxels i and i' within ROI A. Specifically, $S_{ii'}$ is assigned a value of 1 if the Euclidean distance $d_{ii'}$ between voxels i and i' is less than or equal to ε, and 0 otherwise.

We combine the within-region connectivity inference matrix, the underlying graph structures, and spatial contiguity to form a weighted graph $\tilde{\mathcal{G}} = \{\tilde{\mathbf{W}}, \mathcal{S}_A\}$, which serves as the input for our within-region insight detection method referred to as ***SCCN-within***.

2B.2 Dense sub-network extraction

Similar to the bipartite graph G defined in the main text, we decipher the within-region voxel-wise connections using a unipartite graph $H = \left(\cup_c^C H_c\right) \cup H_0$, where each sub-graph $H_c = (R_c, E_c)$ represents an altered sub-network and H_0 is remaining subgraph. Specifically, H_c is a dense subgraph that satisfies (Wu et al., 2022):

$$\sum_{(i,i')\in H_c} \frac{I(\beta_{ii'} \neq 0)}{|E_c|} \gg \sum_{(i,i')\in H_c'} \frac{I(\beta_{ii'} \neq 0)}{|E_c'|}$$

To extract dense sub-graphs H_c from the full graph H, we adopt the objective function with l_0 graph norm shrinkage proposed in Chen et al. (2023):

$$\arg\max_{\cup H_c, C} \log \|\mathbf{U}\|_1 - \lambda_0 \log \|\mathbf{U}\|_0, \tag{.1}$$

where, using our notation, $\|\mathbf{U}\|_1 = \sum_{(i,i')\in H_c} w_{ij} \cdot I(w_{ij} > \gamma)$, $\|\mathbf{U}\|_0 = \sum_{(i,i')\in H_c} I(w_{ij} > \gamma)$, γ is a threshold, and λ_0 is a tuning parameter. Both γ and λ_0 can be estimated using the corresponding likelihood-based method proposed in the main text.

Here, we introduce a customized algorithm to optimize the objective function (.1) by incorporating the spatial constraint. We integrate the inference matrix $\tilde{\mathbf{W}}$ with its spatial-contiguity constraints by performing an element-wise product:

$$\tilde{\mathbf{W}}_{\mathbf{A}} = \tilde{\mathbf{W}} \odot \mathcal{S}_A.$$

As discussed in the main text, $\mathcal{S}_A$ enforces the similarity between non-adjacent voxels to be zero,

effectively removing edges between isolated voxels in the graph H. These spatial-contiguity constraints allow our method to generate outcomes that better adhere to the neurobiological principles emphasizing the coherence of neighboring neuron populations (Thirion et al., 2006). With the input matrix $\tilde{\mathbf{W}}_{\mathbf{A}}$, the algorithm proposed in Chen et al. (2023) becomes applicable for intra-region voxel-level connectivity analysis, with theoretical results provided in that paper to guarantee the consistency and optimality of $\hat{C}, (\hat{H}c)_{c=1,\ldots,\hat{C}}$. We present the implementation tailored to our objective in Algorithm 7 below.

Remarks: The above within region vFC analysis method and our SCCN in the main text jointly contribute to a solution for the $\mathbf{W}_{(n+m)\times(n+m)}$ connectivity matrix.

Algorithm 7 Optimization of function (**??**) for *within-region* subnetwork detection

procedure ALGORITHM (Input: λ and $\tilde{\mathcal{G}} = \{\tilde{\mathbf{W}}, \mathcal{S}_A\}$)
    Calculate $\tilde{\mathbf{W}}_{\mathbf{A}} = \tilde{\mathbf{W}} \odot \mathcal{S}_A$
    for $C = 1, 2, \ldots, n$ **do**
        Ratio-cut spectral clustering $\tilde{\mathbf{W}}_A$ into C networks: $H = \cup_c^C H_c$
        Substitute network sets $\cup H_c$ into objective function (**??**), and obtain the output values
    end for
    return $C, \cup H_c$ that yield the maximum output value
    end function
end procedure

Inference: We employ the network-based inference method proposed in Chen et al. (2023) to statistically test the following hypotheses for each extracted subgraph H_c:

$$\mathbb{H}_0 : H \text{ is a random graph (i.e., no dense unipartite subgraph } H_c \text{ exists)},$$

$$\mathbb{H}_a : \text{At least one dense unipartite subgraph } H_c \text{ exists}.$$

The inference test is conducted using a customized test statistic $T = f(|R|, |H_c|, \gamma_c)$, which reflects the combinatorial probability of a covariate-related sub-network, with $|H_c|$ and γ_c representing the

size and density of clique H_c respectively. Since multiple H_c can be extracted, the family-wise error rate (FWER) control is performed using the permutation procedures proposed in the main text. The procedure provides a corrected p-value for each detected $\hat{H}_c$ (Nichols, 2012).

2B.3 Simulation

In this simulation study, we assessed the proposed method, ***SCCN-within*** by evaluating whether it can extract densely altered sub-area pairs *within* a large brain region. Similar to the main text, we evaluated the performance of our model against comparative methods from two perspectives: (i) multivariate edge-level inference: whether extracted voxel pairs have a high true-positive rate (TPR) and low false-positive rate (FPR); (ii) network-level inference: whether the extracted sub-areas contain maximal true-positive voxels, compared to other unextracted sub-areas.

We generate brain connectome data $\mathcal{Z} = \{\mathbf{Z}^1, \ldots, \mathbf{Z}^S\}$ in a standard two-sample testing setting, with the flexibility to extend it to regression settings if needed. To begin, we generate a unipartite graph $H = \{R, E\}$ consisting of a node set R with size $|R| = 900$ and an edge set E with size $|E| = \binom{900}{2} = 404,550$. Within H, we consider two disease-related sub-networks H_1 and H_2 containing $|R_1| = 90$ and $|R_2| = 130$ nodes, respectively. The true spatial locations of H_1 and H_2 in a simulated 30×30 grid space are presented in Figure A.2 (a1). False negative and false positive noises are added to the ground truth to mimic more realistic neural connectivity, as illustrated in Figure A.2 (a2). Next, we consider two cohorts of subjects with equal sample sizes. For the patient group, we assign connectivity strength values to node-pair $(i, i') \in H$ as follows: for nodes $(i, i') \in H_1 \cup H_2$, we set $z^s_{ii'} \sim \mathcal{N}(\mu_1, \sigma)$; while for all other nodes, we set

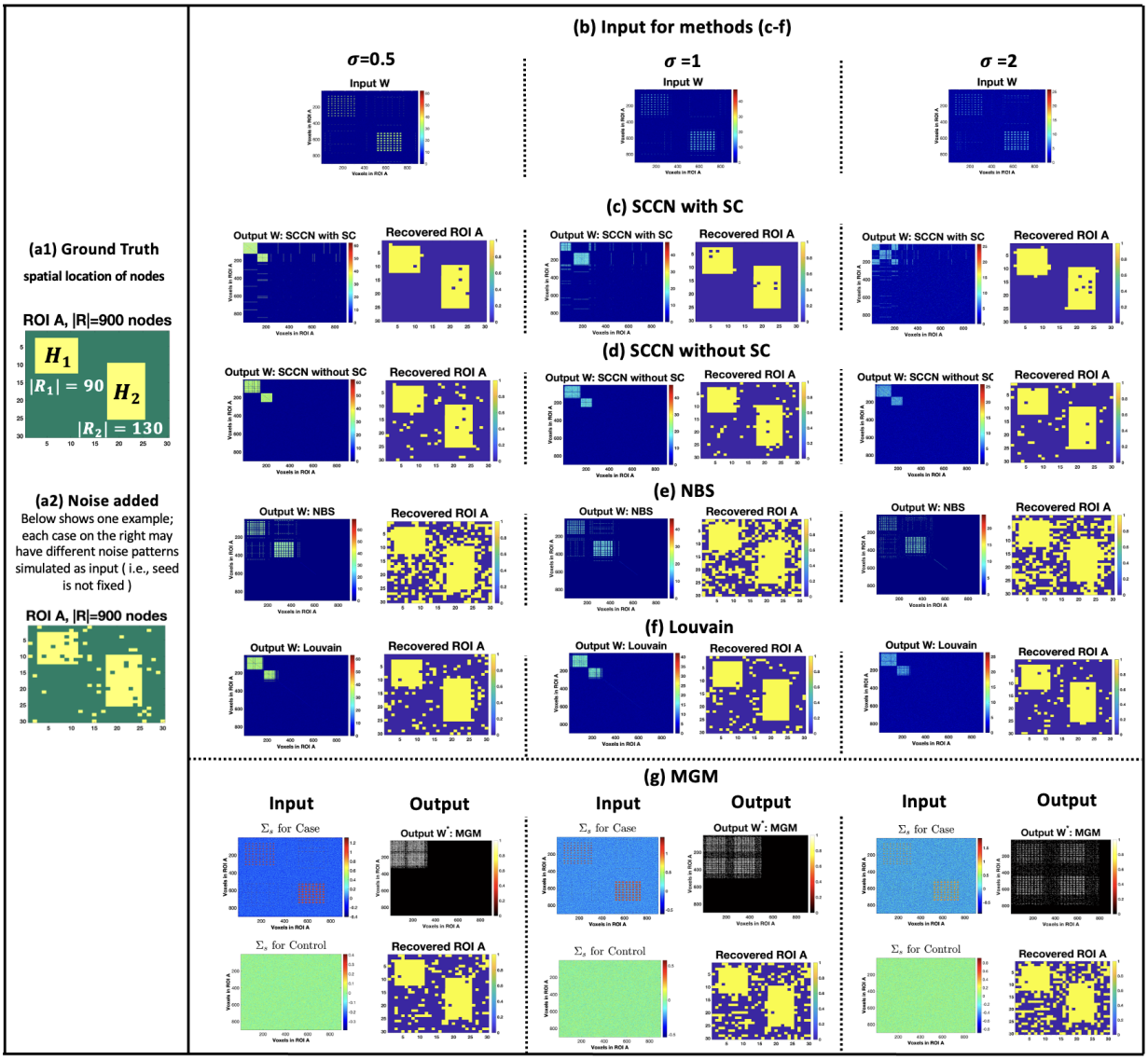


Figure A.2. *A 2D visualization of performance by different network methods using sample size $n = 400$ with balanced case and control subjects. In (a1), the true spatial locations of sub-networks H_1 and H_2 are displayed in a simulated $30 * 30$ grid space. (a2) shows the scenario with false negative and false positive noises added to the ground truth. (b) depicts the heatmap of input inference matrix $\mathbf{W}$ for methods in (c-f), where each element in $\mathbf{W}$ is calculated as $w_{ij} = -\log(p_{ij})$. In (c-f), the reorganized $\mathbf{W}$ is presented, with the disease-related sub-networks detected by each method highlighted at the top. The spatial locations of extracted nodes are also shown. In (c,d), the abbreviation **SC** stands for spatial contiguity. (g) shows the input Σ_s of both case and control group (i.e., the functional connectivity matrix of 900 spatial locations), followed by the output matrix (disease-related nodes are assigned a value of 1; 0 otherwise), and the spatial locations of the nodes of value 1.*

$z^s_{ii'} \sim \mathcal{N}(\mu_0, \sigma)$. As a result, $z^s_{ii'}$ exhibits location-specific characteristics within the ROI. For the control group, we simply set $z^s_{ii'} \sim \mathcal{N}(\mu_0, \sigma)$ for all $(i, i') \in H$. To emulate different effect sizes, or signal-to-noise ratios, we vary the parameters $\theta = \mu_1 - \mu_0$ and σ. Specifically, we set the variance term $\sigma = 0.5, 1, 2$ given $\theta = 1$, resulting in corresponding Cohen's d values of 1.2, 0.8, and 0.5, respectively. Moreover, we consider two different sample sizes $S = 200, 500$, each with balanced controls and patients. All settings with different (σ, S) were simulated 500 times to assess the replicability of the model and the variability of the TPR and FPR.

For each brain connectome data set $\mathcal{Z}$, we conduct edge-wise two-sample tests on $\{z^s_{ii'}\}$ between control and case groups, yielding an inference matrix $\tilde{W}$ by $\tilde{w}_{ii'} = -\log(p_{ii'})$. We then apply *SCCN-within*, with and without integrating spatial-contiguity constraints (i.e., treating $\tilde{\mathbf{W}}$ and $\tilde{\mathbf{W}}_A$ as input, respectively), to compare the performance. We also compare our approach against popular network-based methods for brain connectivity analysis, including NBS (Zalesky et al., 2010) and Louvain (Blondel et al., 2008). Lastly, we apply the Matrix Graph Model (MGM) hypothesis testing method proposed by Xia and Li (2017). Since MGM extracts positive nodes for each individual cohort separately and does not support cross-cohort inference, we adapt our simulation setting as follows: we first obtain the covariance matrix $\Sigma^0_{\mathfrak{s}}$ of 900 spatial locations in ROI A for the control group (i.e., the FC matrix between voxel pairs in our application). We then apply MGM to $\Sigma^0_{\mathfrak{s}}$ to obtain the inference matrix $\mathcal{I}^0$. Specifically, $\mathcal{I}^0$ has the element $\mathcal{I}^0_{ii'} = 1$ if the proposed null hypothesis $H_{0,i,i'} : \omega^0_{\mathfrak{s},i,i'} = 0$ is rejected, where $\omega^0_{\mathfrak{s},i,i'}$ is the (i, i')-th element of the precision matrix $\Omega^0_{\mathfrak{s}}$ of $\Sigma^0_{\mathfrak{s}}$. In contrast, $\mathcal{I}^0_{ii'} = 0$ if $H_{0,i,i'} : \omega^0_{\mathfrak{s},i,i'} = 0$ is not rejected. Following the same procedures, we obtain the covariance matrix $\Sigma^1_{\mathfrak{s}}$ and the inference matrix $\mathcal{I}^1$ for the patient group. Note that both $\Sigma^0_{\mathfrak{s}}$ and $\Sigma^1_{\mathfrak{s}}$ are obtained based on the same spatial and numerical settings defined earlier. Finally, we extract the nodes with different values between $\mathcal{I}^0$ and $\mathcal{I}^1$ as

the disease-related nodes.

Figure A.2(c-g) shows the graph visualization of the detected sub-networks by different methods. Inference results are summarized in Table 1, where node-level True Positive Rate (TPR) and False Positive Rate (FPR) are used as evaluation metrics. We observe that all methods achieve high TPR, but differ in terms of FPR and network-level extraction. Specifically, SCCN with and without integrating spatial contiguity (SC) yield correct network detection and similar TPR; however, SCCN with SC demonstrates a significantly lower FPR, approximately eight times lower. This outcome is expected as SCCN with SC is designed to exclude false positive nodes that are not spatially adjacent. Therefore, in scenarios where true positive nodes gather spatially in a network of interest, SCCN with SC tends to deliver superior results compared to SCCN without SC. Louvain yields very similar TPR and FPR to SCCN without SC. This similarity arises because both methods prioritize community detection by optimizing graph modularity or density. In contrast, while achieving high TPR, NBS shows poor control over FPR (more than 50% false positive nodes when $\sigma \geq 1$). This performance can be attributed to NBS's tendency to extract H_c, which encompasses as many suprathreshold covariate-related edges as possible, leading to an enlarged H_c with low density. Consequently, this also causes NBS to fail in identifying the correct network structures (e.g., NBS recognizes H_1 and H_2 as a single network H_1^*). Lastly, similar to Louvain, MGM outperforms NBS in terms of simultaneously controlling TPR and FPR. MGM ensures high TPR and replicability, as evidenced by its low variance of results across different simulations. MGM generally demonstrates satisfactory control over the FPR while the FPR of MGM is relatively sensitive to the effect size because it does not borrow information from provided spatial constraints. In addition, although MGM identifies most positive voxel pairs in the ROI, these detected voxels cannot automatically form spatially contiguous clusters due to the

lack of spatial information. In summary, the spatial constraints are critical for within-region vFC analysis regarding the TPR, FDR, and network-level detection.

Sample Size: 100 cases v.s. 100 controls				
Methods	**Metric**	σ=0.5	σ=1	σ=2
SCCN (with SC)	TPR	0.998(0.001)	0.998(0.002)	0.996(0.004)
	FPR	0.015(0.021)	0.018(0.025)	0.022(0.027)
	Correct Network Detection	Yes	Yes	Yes
SCCN (without SC)	TPR	0.941(0.030)	0.948(0.017)	0.948(0.045)
	FPR	0.129(0.011)	0.129(0.013)	0.130(0.016)
	Correct Network Detection	Yes	Yes	Yes
Louvain	TPR	0.962(0.011)	0.962(0.012)	0.962(0.012)
	FPR	0.147(0.012)	0.146(0.012)	0.143(0.011)
	Correct Network Detection	Yes	Yes	Yes
NBS	TPR	0.950(0.035)	0.947(0.048)	0.945(0.053)
	FPR	0.402(0.024)	0.590(0.031)	0.594(0.030)
	Correct Network Detection	No	No	No
MGM	TPR	0.951(0.001)	0.951(0.003)	0.948(0.003)
	FPR	0.187(0.011)	0.304(0.013)	0.382(0.013)
	Correct Network Detection	No	No	No

Sample Size: 250 cases v.s. 250 controls				
Methods	**Metric**	σ=0.5	σ=1	σ=2
SCCN (with SC)	TPR	0.998(0.001)	0.998(0.002)	0.996(0.003)
	FPR	0.015(0.019)	0.016(0.022)	0.020(0.023)
	Correct Network Detection	Yes	Yes	Yes
SCCN (without SC)	TPR	0.945(0.030)	0.946(0.017)	0.945(0.035)
	FPR	0.129(0.010)	0.129(0.013)	0.129(0.015)
	Correct Network Detection	Yes	Yes	Yes
Louvain	TPR	0.964(0.009)	0.962(0.011)	0.963(0.011)
	FPR	0.146(0.013)	0.146(0.012)	0.148(0.013)
	Correct Network Detection	Yes	Yes	Yes
NBS	TPR	0.951(0.030)	0.949(0.041)	0.948(0.056)
	FPR	0.427(0.025)	0.572(0.029)	0.576(0.030)
	Correct Network Detection	No	No	No
MGM	TPR	0.952(0.001)	0.951(0.003)	0.948(0.002)
	FPR	0.191(0.010)	0.302(0.011)	0.379(0.011)
	Correct Network Detection	No	No	No

Table 1. *Simulation results. The two sub-tables show the inference results, mean(s.d.), given different sample sizes and variances, where TPR and FPR correspond to the node-level true positive rate and false positive rate. Network detection results indicate whether the method can successfully identify and differentiate the correct disease-related sub-networks H_1 and H_2.*

2B.4 Real data application

Sample characteristic and background

We utilize the fMRI dataset from the schizophrenia study introduced in the main text to perform a real data application of within-region sub-network detection. The dataset comprises 330 individuals, including 148 SZ patients (M/F: 84/64, age: 37.5 ± 14.4) and 182 healthy controls (M/F: 80/102, age: 37.0 ± 16.1). Appendix D.1 provides comprehensive information about participant demographics, the recruitment process, imaging acquisition, and fMRI preprocessing procedures. In this analysis, our specific focus is on the salience network, which is known to

be strongly associated with several core SZ symptoms. Thus, our attention is directed towards the cingulate cortex and left insula, two main components of the salience network, to extract the altered vFC patterns related to schizophrenia within each ROI. According to the Brainnetome Atlas (Fan et al., 2016), the left insula consists of 1762 voxels, while the cingulate cortex comprises 5768 voxels. We applied SCCN-within and the inference test to the within-region inference matrices of the cingulate cortex and insula, $\tilde{\mathbf{W}}^C_{5768\times5768}$ and $\tilde{\mathbf{W}}^I_{1762\times1762}$, respectively, while adjusting for age and sex.

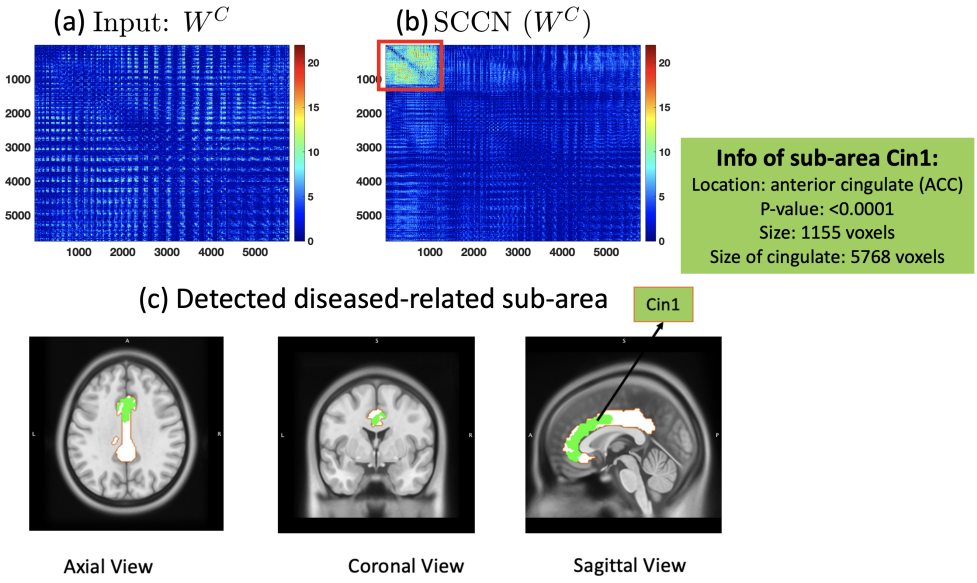


Figure A.4. *(a) A heatmap of the within-Cingulate inference matrix $\tilde{\mathbf{W}}^C_{5768\times5768}$. (b) Results yielded by SCCN: the schizophrenic-altered sub-network is highlighted in red. (c) shows the 3D demonstrations of the detected sub-network from the cingulate cortex, where the within-network nodes are predominantly located in the anterior cingulate cortex (ACC).*

Network-level results Each element in the vFC inference matrices $\tilde{\mathbf{W}}^C$ is $\tilde{\mathbf{W}}^C_{ii'} = -\log(p^C_{ii'})$, where p^C_{ii} is the p-value testing the case-control vFC difference for the node pair (i, i') within the cingulate cortex (Figure A.4(a)). We then perform screening on $\tilde{\mathbf{W}}^C$ using a threshold of $p = 0.005$ determined by the proposed ML-based method. Similar procedures apply to $\tilde{\mathbf{W}}^I$

(Figure A.5(a)). Implementing Algorithm 7 returned a maximum-likelihood estimation (MLE) of $\hat{\lambda}_C = 1.25$ for $\tilde{\mathbf{W}}^C$ and $\hat{\lambda}_R = 1.5$ for $\tilde{\mathbf{W}}^I$. One dense SZ-related sub-network was detected from $\tilde{\mathbf{W}}^C$ and $\tilde{\mathbf{W}}^I$ respectively (marked in red in Figure A.4(b) and Figure A.5(b)), using a significance level of 0.05 from the proposed permutation test in Appendix B.2. 3D demonstrations of the detected sub-network from $\tilde{\mathbf{W}}^C$ and $\tilde{\mathbf{W}}^I$ are shown in Figure A.4(c) and Figure A.5(c), where information regarding the network sizes, p-values, and locations were also specified.

Each element in the vFC inference matrices $\tilde{\mathbf{W}}^C$ is calculated as $\tilde{\mathbf{W}}^C_{ii'} = -\log(p^C_{ii'})$, where $p^C_{ii'}$ is the p-value testing the case-control vFC difference for the node pair (i, i') within the cingulate cortex (see Figure A.4(a)). We then perform screening on $\tilde{\mathbf{W}}^C$ using a threshold of $p = 0.005$ determined by the proposed ML-based method. Similar procedures are applied to $\tilde{\mathbf{W}}^I$ (see Figure A.5(a)). By implementing Algorithm 7, we obtain a maximum-likelihood estimation (MLE) of $\hat{\lambda}_C = 1.25$ for $\tilde{\mathbf{W}}^C$ and $\hat{\lambda}_R = 1.5$ for $\tilde{\mathbf{W}}^I$. One dense SZ-related sub-network is detected from $\tilde{\mathbf{W}}^C$ and $\tilde{\mathbf{W}}^I$ respectively (marked in red in Figure A.4(b) and Table Figure A.5(b)), using a significance level of 0.05 from the proposed permutation test in Appendix B.2. The 3D demonstrations of the detected sub-networks from $\tilde{\mathbf{W}}^C$ and $\tilde{\mathbf{W}}^I$ are shown in Figure A.4(c) and Figure A.5(c), providing information on the network sizes, p-values, and locations.

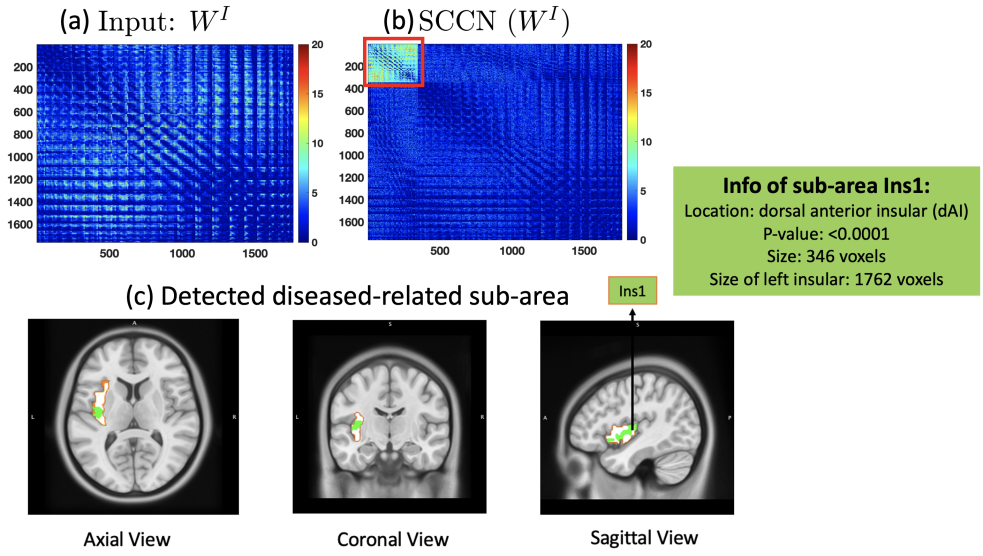


Figure A.5. *(a) A heatmap of the within-Insula inference matrix* $\tilde{\mathbf{W}}^I_{1762\times1762}$. *(b) Results yielded by SCCN: the schizophrenic-altered sub-network is highlighted in red. (c) shows the 3D demonstrations of the detected sub-network from insula, where the within-network nodes are predominantly located in the dorsal anterior insula.*

Biological interpretation of detected sub-areas The voxels in the cingulate cortex that exhibit altered connectivity patterns in schizophrenia, as identified from $\tilde{\mathbf{W}}^C$, are predominantly located in the anterior cingulate cortex (ACC). Similarly, the voxels in the insula that show altered connectivity patterns, as identified from $\tilde{\mathbf{W}}^I$, are concentrated in the dorsal anterior insula. The ACC, known for its involvement in cognitive control processes, has been consistently reported as an abnormal brain region in schizophrenia (Cui et al., 2015; Wang et al., 2015). Additionally, extensive studies have highlighted the important role of the anterior insula in emotional processing, regulation, and cognitive control in schizophrenia (Moran et al., 2013b; Sarpal et al., 2015; Sheffield and Barch, 2016). Our findings obtained through SCCN-within are in line with previous medical research but offer enhanced spatial specificity, providing valuable insights for future treatment approaches.

2C. Proofs and derivations

2C.1 Proof of Lemma 2.1

The proof of Lemma 2 in Chen et al. (2019) for an intra-region stochastic block model can be extended to the inter-region adjacency matrix because of the fact that for a given sub-bipartite graph (U_c, V_d), the number of correct edge assignments to this graph is equal to the product of the number of nodes correctly assigned to U_c and that to V_d. Based on this, we have:

$$\sum_{c=1}^{C}\sum_{d=1}^{D}\left[1-\frac{\left|\,(\hat{S}_c\cap U_c)\bigotimes(\hat{S}_d\cap V_d)\,\right|}{|U_c|\,|V_d|}\right]=\sum_{c=1}^{C}\sum_{d=1}^{D}\left[1-\frac{\left|\,(\hat{S}_c\cap U_c)\,\right|}{|U_c|}\times\frac{\left|\,(\hat{S}_d\cap V_d)\,\right|}{|V_d|}\right]$$

Next, if there exists $x\in(0,1]$ and $y\in(0,1]$, we have $(1-x)(1-y)\geq 0 \leftrightarrow 1-x-y+xy\geq 0 \leftrightarrow 1-xy\leq(1-x)+(1-y)$. Since $\frac{\left|\,(\hat{S}_c\cap U_c)\,\right|}{|U_c|}\in(0,1]$, $\frac{\left|\,(\hat{S}_d\cap V_d))\,\right|}{|V_d|}\in(0,1]$, together with the results of Lemma 2 in Chen et al. (2019), we have

$$\begin{aligned}
&\sum_{c=1}^{C}\sum_{d=1}^{D}\left[1-\frac{\left|\,(\hat{S}_c\cap U_c)\bigotimes(\hat{S}_d\cap V_d)\,\right|}{|U_c|\,|V_d|}\right] &(.2)\\
&=\sum_{c=1}^{C}\sum_{d=1}^{D}\left[1-\frac{\left|\,(\hat{S}_c\cap U_c)\,\right|}{|U_c|}\times\frac{\left|\,(\hat{S}_d\cap V_d)\,\right|}{|V_d|}\right]\\
&\leq\sum_{c=1}^{C}\sum_{d=1}^{D}\left[1-\frac{\left|\,(\hat{S}_c\cap U_c)\,\right|}{|U_c|}\right]+\left[1-\frac{\left|\,(\hat{S}_d\cap V_d)\,\right|}{|V_d|}\right]\\
&\leq D\sum_{c=1}^{C}\left[1-\frac{\left|\,(\hat{S}_c\cap U_c)\,\right|}{|U_c|}\right]+C\sum_{d=1}^{D}\left[1-\frac{\left|\,(\hat{S}_d\cap V_d)\,\right|}{|V_d|}\right]\\
&\leq\tau_A^{-1}(2+\varepsilon_A)\frac{ndCD}{\Lambda_A^2}+\tau_B^{-1}(2+\varepsilon_B)\frac{mdCD}{\Lambda_B^2}, &(.3)
\end{aligned}$$

2C.2. Proof of Theorem 2.2

Theorem 1 in Chen et al. (2019) states that, by maximizing its objective function via ratio-cut partitioning iteratively, the number of mis-assigned nodes N_{node} from a graph $G = (V, E)$ satisfies $N_{\text{node}} = o_p(n_{\min})$ as $n \to \infty$, where $n_{\min}$ is the size of the smallest subgraph in G. Therefore, let $n_{\min}, m_{\min}$ be the sizes of the smallest possible subgraphs in U and V, respectively, we have

$$\lim_{|U|\to\infty,|V|\to\infty} \frac{|U||V|}{n_{\min}m_{\min}} = \lim_{|U|\to\infty} \frac{|U|}{n_{\min}} \lim_{|V|\to\infty} \frac{|V|}{m_{\min}} = 0$$

Consequently, $|U||V| = o_p(n_{\min} * m_{\min})$.

2C.3. Proof of Theorem 2.3

The connectivity inference matrix $\mathbf{W}$ is usually obtained from a finite dataset. When the cluster number C is pre-determined for region A, **Theorem 1** in Luxburg et al. (2004) indicates that the partitions constructed by ratio-cut spectral clustering "from the first r eigenvectors on the finite samples converge almost surely to a limit partition of the whole data space". Similarly, clustering on region B with a cluster number D also yields a convergent estimate. In addition, during each iterative search for the number of clusters C, D, the objective function (1) keeps being updated and eventually finds the maximizers $\hat{C}, \hat{D}$. Finally, by Theorem 1, ratio-cut spectral clustering on $\mathbf{W}_A, \mathbf{W}_B$ using $\hat{C}, \hat{D}$ will converge almost surely to the true community structure where false-positive edge assignments are negligible in very large bipartite graphs.

2C.4. Construction of the MDL-based test statistics

Consider families of models $\{\mathcal{M}_\gamma : \gamma \in \Gamma\}$, where each $\mathcal{M}_r = \{q_\theta : \theta \in \mathbf{\Theta}\}$ is a statistical model used to model data $x^n := (x_1, \ldots, x_n)$ with each $x_i \in \mathbf{X}$, for some input space $\mathbf{X}$. MDL associates each $\mathcal{M}_r$ with a universal distribution $\bar{q}_\theta$, and the indices $\Gamma := \{1, 2, \ldots, \gamma_{max}\}$ with a distribution, say π. The best explanation of the given data x^n is provided by the model $\mathcal{M}_r$ minimizing the total codelength:

$$q_\gamma(x^n) = \min_{\gamma \in \Gamma} \{L_\pi(\gamma) + L_{q_\gamma}(x^n)\}, \tag{.4}$$

where $L_\pi(\gamma) := -\log \pi(\gamma)$ measures the codelength (e.g. the number of bits) needed to encode the model using a particular code, and $L_{q_\gamma}(x^n) := -\log q_\gamma(x^n)$ measures the codelength needed to encode the data x^n. In our case, the input data is an edge-wise statistical inference matrix $\mathbf{W}_{n \times m}$. Let $\mathbf{X}$ be standardized $\mathbf{W}$ such that elements $x_{ij} \in \mathbf{X}$ have zero mean and unit variance. Let $U_c \bigotimes V_d$ be a bicluster generated by performing Algorithm1 on $\mathbf{X}$. Then the codelength describing the data matrix $\mathbf{X}$ comes from both the bicluster $U_c \bigotimes V_d$ and the residual matrix.

First, for the yielded submatrix $U_c \bigotimes V_d$, we consider three elements to be described: location, size, and magnitude of entries inside. With slight abuse of notation, we let $|U_c| = c$ and $|V_d| = d$. Then the codelength needed to describe the location of the submatrix is $\log_2[\binom{n}{c}\binom{m}{d}]$ bits since a total of $\binom{n}{c}\binom{m}{d}$ different submatrices with size $c \times d$ can be selected from $\mathbf{X}_{n \times m}$. Also, coding the size of the submatrix $U_c \bigotimes V_d$ requires $\log_2(nm)$ bits since $c \in 1, \ldots, n$ and $d \in 1, \ldots, m$. Lastly, to code the submatrix magnitude, since $\mathbf{X} = standardized(\mathbf{W})$,where $\mathbf{W}$ consists of post-screened $-\log(p_{ij})$, we assume each elements x_{ij} lies within the range $[0, 30]$

(usually a value of $-\log(p_{ij}) > 26$ is rarely observed). Therefore, the codelength to describe the submatrix magnitude would be approximately $\log_2 30 + C$, where C is the precision of digits after the decimal point.

Next, we consider the residual matrix $\mathbf{R}$. The total variation from the full data set $\mathbf{X}$ is $\sum_{i \notin U_c} \sum_{j \notin V_d} (x_{ij} - \mu_0)^2 = mn$ as $mean(\mathbf{X}) = \mu_0 = 0$, and $var(\mathbf{X}) = 1$. The variation explained by the submatrix $U_c \bigotimes V_d$ is approximately $\sum_{i \in U_c} \sum_{j \in V_d} (\mu_1 - \mu_0)^2 = cd\mu_1^2$, where μ_1 is the mean of the submatrix. Hence, the residual matrix $\mathbf{R}$ explains a total of $mn - cd\mu_1^2$ variation, and therefore by the Central Limit Theorem, entries in $\mathbf{R}$ approximately follows a $N(0, \frac{nm - cd\mu_1^2}{nm - cd})$ distribution. In information theory, for a random variable Y with probability density function $f(y)$, it typically requires $-\log_2(f(Y)) + C$ bits to code it. Thus, for a standard normal random variable $Z \sim N(0, 1)$, the expected codelength is $L_Z = -\int f(z) \log_2 \Big(f(z)\Big) dz + C$ bits, and for a random variable $Y \sim N(0, \sigma^2)$, its expected codelength is:

$$
\begin{aligned}
L_Y &= -\int \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{y^2}{2\sigma^2}}\ \log_2\Big(\frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{y^2}{2\sigma^2}}\Big) dy \\
&= -\int f(z) \log_2 \Big(f(z)\frac{1}{\sigma}\Big) dz \\
&= L_Z - \log_2(\frac{1}{\sigma}) \int f(z) dz \\
&= L_Z + \frac{1}{2}\log_2(\sigma^2)
\end{aligned}
$$

Therefore, coding each element in $\mathbf{R}$ requires $L_Z + \frac{1}{2}\log_2(\frac{nm - cd\mu_1^2}{nm - cd})$ bits on average. Combining the codelengths needed for both $U_c \bigotimes V_d$ and $\mathbf{R}$ gives the minimum description

length for explaining the input data set $\mathbf{X}$ using a $c \times d$ bicluster $U_c \bigotimes V_d$:

$$
\begin{aligned}
mdl(U_c \bigotimes V_d) &= \log_2 \left[\binom{n}{c} \binom{m}{d} \right] + \log_2(nm) + \log_2 30 + C \\
&\quad + (nm - cd)\Big(L_Z + \frac{1}{2}\log_2(\frac{nm - cd\mu_1^2}{nm - cd})\Big) \\
&\approx const + \log_2 \left[\binom{n}{c} \binom{m}{d} \right] + \Big(\frac{1 - \mu_1^2}{2 \ln 2} - L_Z\Big)cd,
\end{aligned}
$$

The approximation of the second line above comes from the first order Taylor polynomial near zero: $\log_2(x + 1) = \frac{x}{\ln 2} + o(x)$. For each bicluster $U_c \bigotimes V_d$, there is an associated $mdl(U_c \bigotimes V_d)$ value.

2D. Additional information on schizophrenia data analysis

2D.1. fMRI data acquisition and pre-processing procedures

All 330 participants (148 patients with schizophrenia and 182 healthy controls, demographic information provided in Table A.1 provided written informed consent as approved by the local Institutional Review Board. The participants were required for a large ongoing study of the effects of cerebral disconnectivity in schizophrenia. Diagnoses were confirmed with the Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders-IV (DSM-IV) in all participants. Major medical and neurological illnesses, history of head injury with cognitive sequelae, and diagnosis of intellectual disability were exclusionary. Control participants had no past or present DSM-IV Axis I psychiatric disorder and no family history of psychosis in the prior two generations. All subjects were assessed for research studies at the local research center using uniform recruitment criteria. The neurological and clinical assessments as well as the inclusion of

a control group were kept the same.

	Mean age (SD)	Age range	Gender (M/F)	Age-of-onset	Illness duration
SCZ (N=148)	37.5(14.4)	18.63	84/64	18.1 ± 8.5	21.5 ± 15.1
HC (N=182)	37(16.1)	16.54	80/102	NA	NA

Table A.1. Subject Demographic Information. Values for age are in years.

Imaging data were collected using a Siemens 3 T TRIO MRI (Erlangen, Germany), running VB17 software and equipped with a 32-channel RF head coil at the local research center. Resting-state functional T2*-weighted images were obtained using a single-shot gradient-recalled, EPI pulse sequence (TR: 2000 ms, TE: 30 ms, 128×128 matrix, 1.72×1.72 mm^2 in-plane resolution, 4 mm slice thickness, 37 axial slices, and 444 volumes). Following the previously published procedures, data were preprocessed in AFNI and MATLAB (MathWorks, Inc., Natick, MA). Nuisance variables such as the linear trend, 6 motion parameters, their 6 temporal derivatives, and time courses from the local white matter and cerebrospinal fluid from lateral ventricles were modeled using multiple linear regression analyses, which were then removed as regressors of no interest. Time points with excessive motion (> 0.2 mm), estimated as the magnitude of the displacement from one time point to the next, including neighboring time points and outlier voxels fraction (> 0.1) were censored from statistical analysis. Images were spatially normalized to the ENIGMA EPI template in Montreal Neurological Institute standard space for group analysis.

2D.2. Salience network disrupted connectivity

In this study, we use the Brainnetome Atlas to depict the nodes in the brain connectome graph. The Brainnetome Atlas is one of the commonly used brain atlases which parcellates the brain into 246 connectivity-based sub-regions. Time series are extracted from subjects' insular (labeled

as 163-174 in Brainnetome Atlas) and cingulate cortex (labeled as 175-188 in the Brainnetome Atlas). let RI_i be the i-th altered sub-area detected from the right insula that is connected to the j-th sub-area, CR_j, from the cingulate cortex. Also, let N_{L_i}, N_{R_i}, $N_{C_{L_j}}$, $N_{C_{R_j}}$ denote the number of voxels in the corresponding sub-area. Based on the Brainnetome Atlas, there are 1762 voxels in left insular, 1442 voxels in the right insular, and 5768 voxels in the cingulate cortex.

In $\mathbf{W}^L$, the total number of abnormal edges detected by SCCN is 840442, which accounts for approximately 8.27% of all edges connected between the two regions. Among the nine detected pairs of sub-areas from $\mathbf{W}^L$, there are four sub-areas identified from left insular: $N_{L_1} = 51, N_{L_2} = 32, N_{L_3} = 297, N_{L_4} = 260$, and five sub-areas identified from cingulate cortex: $N_{C_{L_1}} = 266, N_{C_{L_2}} = 116, N_{C_{L_3}} = 282, N_{C_{L_4}} = 975, N_{C_{L_5}} = 422$ (see Figure 4(e) for the network connectome between L_j and C_{L_j}). Specifically, L_1 and L_2 are located in the dorsal anterior insula (dAI) region while L_3 and L_4 are located in the dorsal posterior insular region. C_{L_1}, C_{L_2}, and C_{L_3} are located in the pregenual anterior cingulate cortex (pACC) region, while C_{L_4} and C_{L_5} are located near the anterior midcingulate cortex (aMCC) and midcingulate cortex (MCC). A 3D demonstration of detected results from $\mathbf{W}^L$ is provided in Figure 4 in the main body of the paper.

In $\mathbf{W}^R$, the total number of abnormal edges found is 101241, which accounts for approximately 1.22% of all edges connected between the two regions. Among the ten detected pairs of sub-areas from $\mathbf{W}^R$, there are seven sub-areas identified from right insular: $N_{R_1} = 36, N_{R_2} = 97, N_{R_3} = 9, N_{R_4} = 68, N_{R_5} = 40, N_{R_6} = 88, N_{R_7} = 45$, and nine sub-areas identified from cingulate cortex: $N_{C_{R_1}} = 116, N_{C_{R_2}} = 90, N_{C_{R_3}} = 129, N_{C_{R_4}} = 113, N_{C_{R_5}} = 220, N_{C_{R_6}} = 288, N_{C_{R_7}} = 152, N_{C_{R_8}} = 251, N_{C_{R_9}} = 70$ (see Figure 4(e) for the network connection structures between R_i and C_{R_j}). Specifically, R_1 and R_2 are located in the dorsal anterior insula (dAI) region, R_4 and R_5

are located in the dorsal posterior insular region, while R_3 sits somewhere in the middle between the two regions. C_{R_1} - C_{R_5} are located in the pregenual anterior cingulate cortex (pACC) region, while C_{R_6} - C_{R_9} are located near the anterior midcingulate cortex (aMCC) and midcingulate cortex (MCC). A 3D demonstration of detected results from $\mathbf{W}^R$ is shown in Figure A.6.

2D.3. Temporal-thalamic disrupted connectivity

We apply SCCN to investigate the localized altered connectivity between the middle temporal gyrus on the right hemisphere and the bilateral thalamus in SZ patients. Implementing Algorithm 2 returned the maximum likelihood estimation of $\hat{\lambda} = 0.625$ for both $\mathbf{W}^{(\mathrm{Tem_{right}},\mathrm{Tha_{left}})}$ and $\mathbf{W}^{(\mathrm{Tem_{right}},\mathrm{Tha_{right}})}$. Given the estimated $\hat{\lambda}$'s, Algorithm 1 returned the number of clusters $\hat{C}_L = 426, \hat{D}_L = 75$ for $\mathbf{W}^{(\mathrm{Tem_{right}},\mathrm{Tha_{left}})}$, and $\hat{C}_R = 65, \hat{D}_R = 72$ for $\mathbf{W}^{(\mathrm{Tem_{right}},\mathrm{Tha_{right}})}$. The MDL-based test returned 11 abnormal sub-area pairs for both $\mathbf{W}^{(\mathrm{Tem_{right}},\mathrm{Tha_{left}})}$ and $\mathbf{W}^{(\mathrm{Tem_{right}},\mathrm{Tha_{right}})}$ (marked in red in Figure A.7(2) and Figure A.8(2). 3D demonstrations of the corresponding detected results are also provided in these two figures (using a significance level of 0.05 from the MDL-based permutation test). All extracted sub-area pairs show well-organized topological structures. In summary, the aberrant vFC patterns from $\mathbf{W}^{(\mathrm{Tem_{right}},\mathrm{Tha_{left}})}$ and $\mathbf{W}^{(\mathrm{Tem_{right}},\mathrm{Tha_{right}})}$ are gathered between the anterior/medial part of the middle temporal gyrus on the right hemisphere and medial bilateral thalamus. These results in general align well with previous medical findings: temporal-thalamic dysconnectivity serves as a core neurological feature of SZ that underpins positive symptoms(Ferri et al., 2018). However, our new results provide much more spatially specific information on SZ, which may be helpful for future SZ treatments and studies.

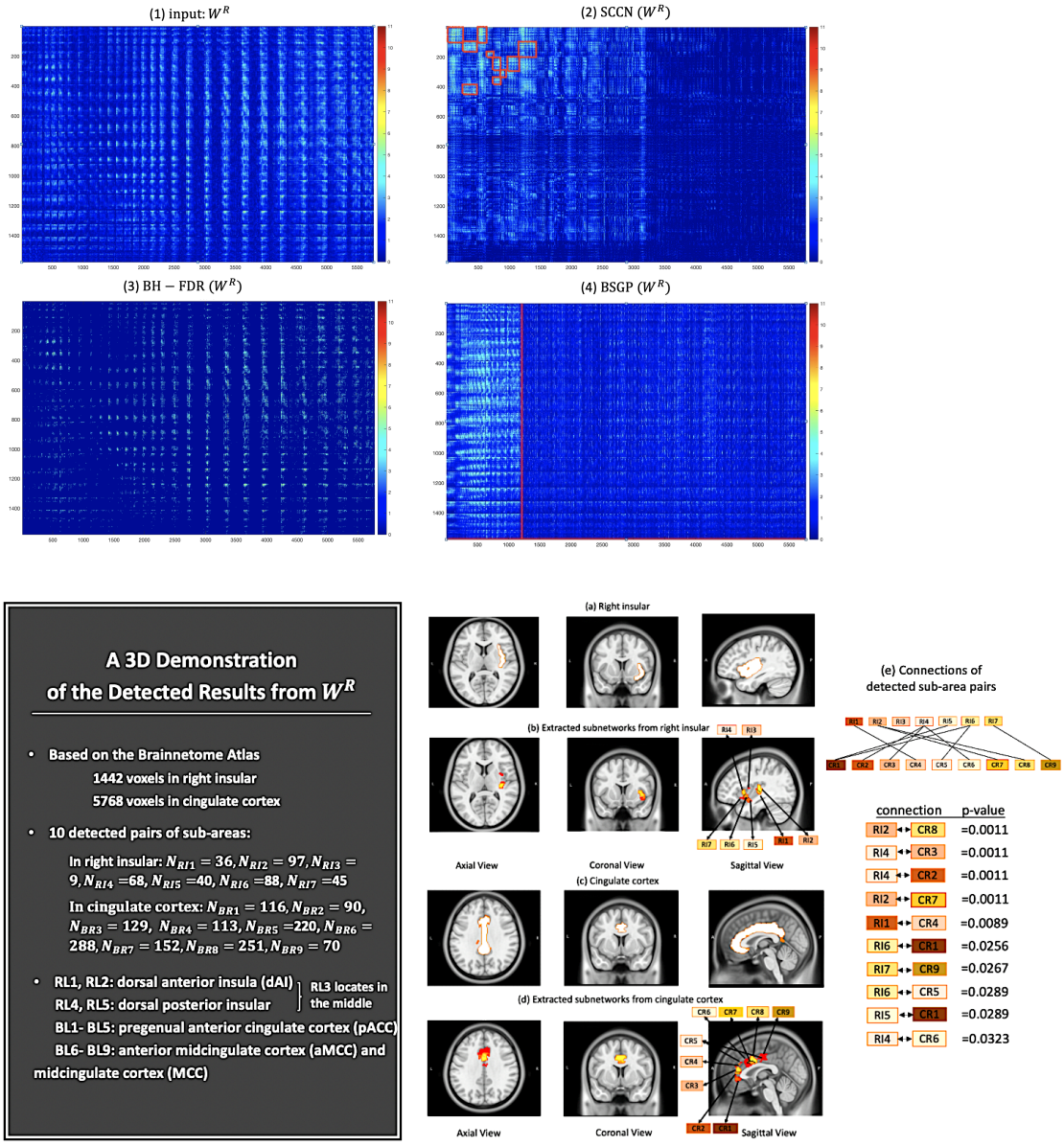


Figure A.6. *(1) A heatmap of* $\mathbf{W}^R$*: rows and columns correspond to the voxels from the right insula and the cingulate cortex, respectively. A hotter entry indicates a more differentially expressed voxel pair between clinical groups adjusted for other covariates. (2) Results yielded by SCCN: positive sub-area pairs that pass the MDL-based permutation test are highlighted in red boxes. There are many edges with small p-values outside the red boxes (e.g., in the bottom left corner) because they are not spatially contiguous to those inside the boxes, and are automatically excluded by SCCN. (3) Results yielded by BH-FDR: with* $q = 0.05$*, no sub-area pairs were detected. (4) Results yielded by BSGP: only one informative yet much less dense sub-area pair was detected. The detected sub-area pair was also lack of spatial contiguity and specificity. (a)-(e) show a 3D demonstration of the detected results from* $\mathbf{W}^{(\mathrm{BG}_{\mathrm{left}}, \mathrm{Ins}_{\mathrm{left}})}$.

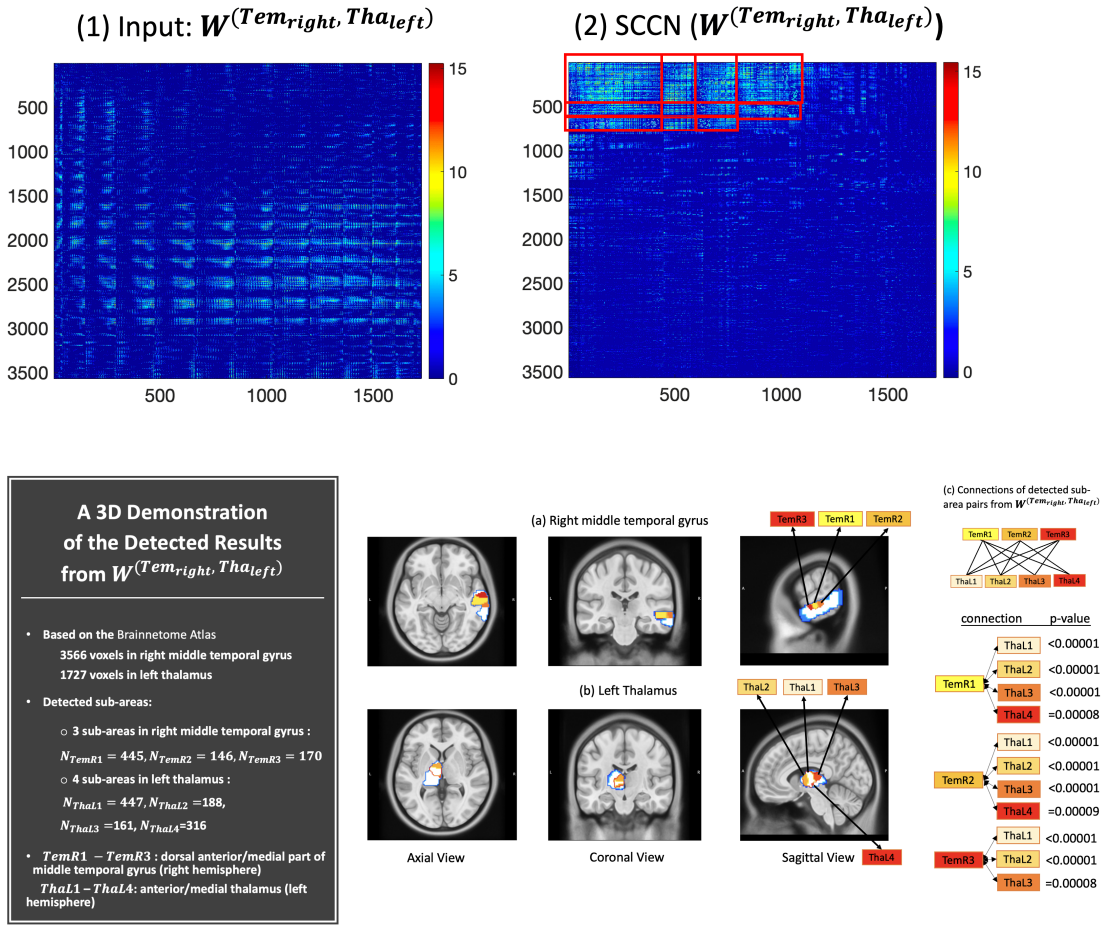


Figure A.7. *(1) A heatmap of* $\mathbf{W}^{(\mathrm{Tem}_{\mathrm{right}}, \mathrm{Tha}_{\mathrm{left}})}$*: rows and columns correspond to voxels from the middle temporal gyrus on the right hemisphere and the left thalamus, respectively. (2) Results yielded by SCCN: altered sub-area areas that pass the MDL-based permutation test are highlighted in red boxes. (a)-(c) shows the 3D demonstration of the 11 detected altered sub-areas from* $\mathbf{W}^{(\mathrm{Tem}_{\mathrm{right}}, \mathrm{Tha}_{\mathrm{left}})}$*. (a)(b) show a 3D demonstration of the detected results from* $\mathbf{W}^{(\mathrm{Tem}_{\mathrm{right}}, \mathrm{Tha}_{\mathrm{left}})}$*, where areas contained in blue borders are the original ROIs. Based on the p-values shown in (e), most positive sub-area pairs are located in the dorsal anterior/medial part of the middle temporal gyrus on the right hemisphere and anterior/medial thalamus on the left hemisphere.*

2E. Additional information on UK Biobank smoking data analysis

2E.1. Subject selection

We selected our study participants from the first imaging visit. Out of the total of 40,092 subjects who provided viable fMRI data, we classified those as nicotine-addicted who met the

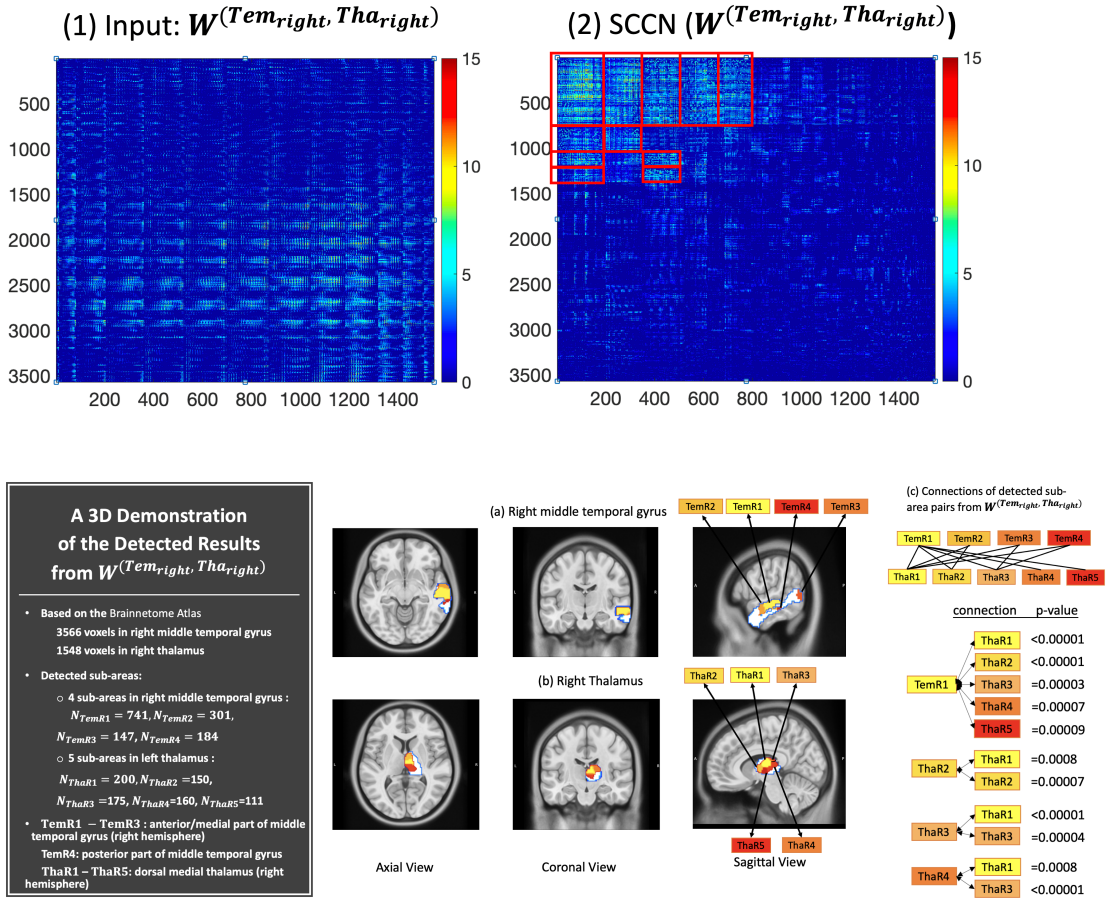


Figure A.8. *(1) A heatmap of* $\mathbf{W}^{(\mathrm{Tem}_{\mathrm{right}}, \mathrm{Tha}_{\mathrm{right}})}$*: rows and columns correspond to voxels from the middle temporal gyrus on the right hemisphere and the left thalamus, respectively. (2) Results yielded by SCCN: altered sub-area areas that pass the MDL-based permutation test are highlighted in red boxes. (a)-(c) shows the 3D demonstration of the 11 detected altered sub-areas from* $\mathbf{W}^{(\mathrm{Tem}_{\mathrm{right}}, \mathrm{Tha}_{\mathrm{right}})}$*. (a)(b) show a 3D demonstration of the detected results from* $\mathbf{W}^{(\mathrm{Tem}_{\mathrm{right}}, \mathrm{Tha}_{\mathrm{right}})}$*, where areas contained in blue borders are the original ROIs. Based on the p-values shown in (e), most positive sub-area pairs are located in the dorsal anterior/medial part of the middle temporal gyrus on the right hemisphere and anterior/medial thalamus on the left hemisphere.*

following criteria: 1. They were currently smoking, as per data field 1239 ("Current tobacco smoking"); 2. they consumed ten or more cigarettes per day (CPD), according to data field 3456 ("Number of cigarettes currently smoked daily"). These criteria resulted in the identification of 1,353 nicotine-addicted subjects. To establish a control group, we collated data from fields 1239 ("Current tobacco smoking"), 1249 ("Past tobacco smoking"), 2644 ("Light smokers, at least 100 smokes in lifetime"), and 20116 ("Smoking status"). We included subjects who had a history of tobacco exposure but were not currently smoking. Specifically, these subjects had a previous smoking history that included occasional smoking and at least 100 lifetime smokes. This group was designated as non-nicotine addicted. As a result, we identified 1,916 subjects who formed the control group. All relevant information from different data fields listed above can be found at https://biobank.ctsu.ox.ac.uk/crystal/search.cgi. In summary, we included a total of $1353 + 1916 = 3269$ subjects in the study.

2E.2. fMRI data acquisition and pre-processing procedures

To obtain the fMRI data with quality controls, we first acquired the pre-processed rs-fMRI data in the file named "filtered_func_data_clean" for each subject provided by UK Biobank. UK biobank uses identical scanner hardware and software in all sites (3T Siemens Skyra, 32-channel Siemens head RF coil, software VD13) This data underwent a series of processing steps outlined in Section 3.11, page 18 of the Brain Imaging Documentation: "1. motion correction using MCFLIRT (Jenkinson et al., 2002); 2. grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor; 3. highpass temporal filtering (Gaussian-weighted least-squares straight line fitting, with $\sigma = 50.0$s); 4. EPI unwarping (utilizing the field maps as described above); 5. GDC

unwarping; 6. Finally, structured artifacts are removed by ICA+FIX processing (Independent Component Analysis followed by FMRIB's ICA-based X-noisier (Beckmann and Smith, 2004; Salimi-Khorshidi et al., 2014)).” Following these steps, the pre-processed data was obtained with controlled quality (Alfaro-Almagro et al., 2018): leave-one-out testing demonstrated a (mean/median) classification accuracy of 99.1/100.0% for non-artifact components and 98.1/98.3% for artifact components. We next applied a transformation to map functional data into a standard space (MNI 152) and applied the brain mask to the transformed data. Finally, we extracted times series from each voxel of interest in the masked data to calculate the voxel-wise functional connectivity.

2E.3. Covariates and Confounders

UK biobank uses identical scanner hardware and software across all sites (3T Siemens Skyra, 32-channel Siemens head RF coil, software VD13), so there was no need for adjustments to the scanner or acquisition protocol parameters. In our work, we carefully accounted for a conventional set of confounds commonly employed in typical imaging studies. Similar sets of confounds have been used and described in the UK Biobank and Enigma Projects (Miller et al., 2016) and have been acknowledged as ”conventional simple” confounds in Alfaro-Almagro et al. (2021). These confounds encompassed age, sex, head size, site, and head motion. Additionally, we incorporated educational level and Body Mass Index (BMI) as they are known to have significant associations with brain functional connectivity strength (Arenaza-Urquijo et al., 2013; Kullmann et al., 2012). Specifically, given that head size and motion estimates were corrected during the pre-processing steps, we did not treat them as covariates in the subsequent regression analysis. In summary, we

focused on healthy participants in the nicotine-addiction study and considered their age, sex, site, educational level, and BMI as nuisance covariates. The primary covariate of interest pertains to the distinction between current smokers (individuals who currently smoke more than ten cigarettes per day) and previous light smokers (individuals who have only tried a few cigarettes in the past but are not currently addicted to nicotine products). We obtained the inference matrix $\mathbf{W}$ with each element $w_{ij} = -\log(p_{ij})$ for each analysis on considered brain region pairs by regressing participants' FCs on the primary covariate of interest while adjusting for the forenamed nuisance covariates.

2E.4. Network detection results

In this section, we present the network detection results of the remaining 11 cases for the nicotine addiction research data, where all other settings stay the same (e.g., datasets, clinical background, analytical approach, etc.). It would be helpful to restate the abbreviation of the selected regions: basal ganglia (BG), hippocampus(Hippo), and insular gyrus (Ins). Again, we label these regions based on the 246 Brainnetome Atlas: there are 2345 voxels in left BG, 2375 voxels in right BG, 1162 voxels in left Hippo, 1099 voxels in right Hippo, 1762 voxels in left Ins, and 1577 voxels in the right Ins. We followed the same computational procedures as presented in the main manuscript and obtained the edge-wise connectivity inference matrices for the 11 ROI pairs across clinical groups. We then applied SCCN and the MDL-based permutation test to these inference matrices in order to extract densely altered sub-area pairs with spatial-contiguity constraints. By plotting the detected results on a 3D brain space, we have some new findings:

(1)The (BG, Ins) pair: similar to the results of $\mathbf{W}_{2345\times1762}^{(\mathrm{BG_{left}},\mathrm{Ins_{left}})}$ presented in the paper, we

found that the altered vFC gather between the medial interior part of right BG to the bilateral Ins as well. In addition, the subarea pairs (bilateral interior BG, posterior Ins) and (putamen (i.e., the outermost portion of BG), interior Ins) also densely contain altered voxel pairs.

(2)The (BG, Hippo) pair: It is interesting to see that almost the entire region of Hippo and BG exhibit differentially expressed connections between current smokers and previous light smokers. SCCN found no specific sub-area pairs between the two large ROIs that primarily drive the connectivity difference between case and control. This aligns with previous findings of Wetherill et al. (2014) that there was increased connectivity in smokers in the whole hippocampus.

(3)The (Hippo, Ins) pair: After applying SCCN, we found that the subarea pairs (bilateral Hippo, anterior Ins) and (parahippocampus, posterior Ins) densely contain altered voxel pairs between current smokers and previous light smokers. Specifically, voxel-pair connectivity increased in smokers within the (bilateral Hippo, anterior Ins) pair and decreased in smokers within the (parahippocampus, posterior Ins) pair. These results also align with previous discoveries of Moran et al. (2013a) and Wetherill et al. (2014).

2F. Additional information on negative control analysis

We implemented SCCN on the inference matrix $\mathbf{W^0}$ of the negative control data with different values of variance σ and sample size S, where $S = 1000$ and $\sigma = 1$. Figure A.9 shows a detection result from $1/1000$ simulation, where no sub-area pairs were detected, and only 202 positive singletons (i.e., edges connecting region A and B) were found out of $900 \times 1600 \approx 1.5$ millions edges. Combing all 1000 iterations, the subarea-level false positive findings were 0, and the edge-wise FPR (supra-threshold voxel-pairs) is 6.82×10^{-5} $(\text{std. } 1.29 \times 10^{-5})$, which with

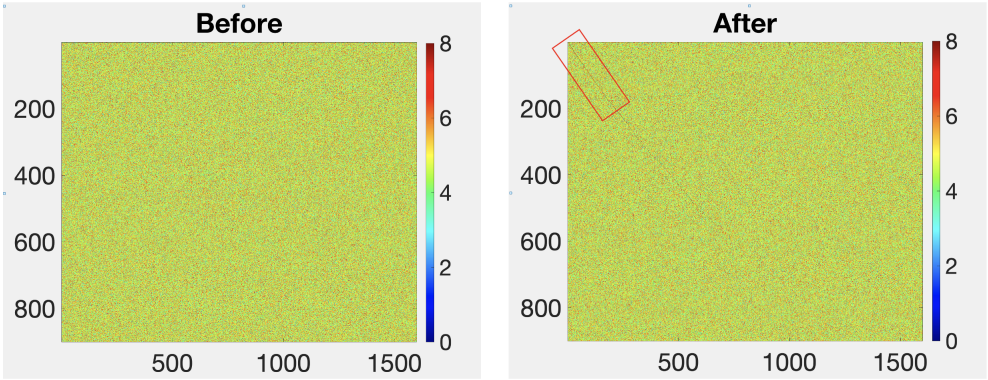


Figure A.9. *Results of negative control analysis. The left shows the heatmap of the inference matrix* $\mathbf{W}$ *obtained with* $\sigma = 1$ *and* $S = 1000$*, when there is no significant connectivity difference between patients and controls. The right shows the heatmap of reshuffled voxel-pair inference* $(-\log p_{ij}, i \in Region\ A, j \in Region\ B)$ *obtained after implementing SCCN. No abnormal sub-area pairs are detected. Different sample sizes and variances yield very similar results.*

consistent with the pre-determined alpha level ($\mathbb{E}(p) = 0.00005$).

Appendix : HIMA

3A. Additional information on real imaging data

Imaging was obtained using a Siemens 3T PRISMA scanner with a 64-channel head coil. Imaging parameters were: field-of-view=224mm; matrix size=64x64; 12 axial slices (10% oversampling); resolution = 3.5x3.5x8mm; turbo factor=14; EPI factor=64; bandwidth=3,25Hz/pixel; TE=36.5ms; TR=4,000ms; label/control duration=1,500ms; centric ordering; and timing of background suppression was optimized. A two-stage approach was utilized21. Fifteen repetitions were acquired during the flow encoding arterial spin tagging scan at the post-labeling delay (PLD)=900ms for estimating arterial transit time (ATT). Kw was calculated from scans at PLD=1,800ms with b=0 and 50s/mm2, with twenty repetitions acquired per b=value. Kw measured six-weeks apart has shown good reproducibility. Cerebral blood flow (mL/100g/min) was quantified from signals acquired at PLD=1,800ms without diffusion preparation.

Images were corrected for head motion and Kw maps were calculated by a total-generalized-variation regularized single-pass-approximation model using the tissue/capillary fraction of ASL signal at PLD 1,800ms, incorporating ATT, arterial blood T1 (1,660ms), and brain tissue as inputs18. Voxel-wise T1 map was fitted from background suppressed control images. Data were processed using FSL Bayesian Inference for ASL pipeline after partial volume correction. Equilibrium magnetization (M0) images were collected using a TR=10s. Kw represents the ratio

between capillary permeability surface area product of water (PSw) and the distribution volume of water tracer in capillary space (Vc) (PSw/Vc) to estimate the putative water exchange rate across the BBB.

3B. Theoretical justifications of HIMA

In this section, we present the asymptotic properties of our procedure. The adopted estimator for the posterior mode $\boldsymbol{\Sigma}^{\text{mode}}$ has shown to converge in distribution to the true estimator $\boldsymbol{\Sigma}_0$ in the likelihood $p(\boldsymbol{Y}|\boldsymbol{\Sigma}_0, \mu_0)$ when the dimension p is large. Generating $\boldsymbol{\Sigma}$ based on it the posterior mode in the relaxed posterior P-step shall yield reliable imputations. Here is the heuristic justification for this asymptotic result. The estimator for the posterior mode $\boldsymbol{\Sigma}^{\text{mode}}$ is obtained by minimizing the expected estimation loss $L(\boldsymbol{\Sigma}_0, \hat{\boldsymbol{\Sigma}})$, which is a regular, asymptotically linear (RAL) estimator. Brain imaging data typically has a large number of voxels p, and thus a large number of observed voxels p_{obs} ordinarily. Tsiatis (2006) established that with an RAL estimator $\hat{\boldsymbol{\Sigma}}$,

$$p_{\text{obs}}^{1/2}\left(\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}_0\right) \approx p_{\text{obs}}^{-1/2} \sum_{i=1}^{N} \varphi\left(\mathbf{R}_i, \mathbf{Z}_{(\mathbf{R}_i)}\right),$$

where $\varphi\left(\mathbf{R}, \mathbf{Z}_{(\mathbf{R})}\right)$ is the influence function of the estimator $\hat{\boldsymbol{\Sigma}}$, which has mean zero. Based on Tsiatis's work, we can obtain the asymptotic variance of $\hat{\boldsymbol{\Sigma}}$:

$$\text{var}\left\{\varphi\left(\mathbf{R}, \mathbf{Z}_{(\mathbf{R})}\right)\right\} = E\left\{\varphi\left(\mathbf{R}, \mathbf{Z}_{(\mathbf{R})}\right)\varphi\left(\mathbf{R}, \mathbf{Z}_{(\mathbf{R})}\right)^T\right\} := \boldsymbol{\zeta}.$$

It thus follows that the convergence $\hat{\boldsymbol{\Sigma}} \xrightarrow{d} \boldsymbol{\Sigma}_0$ in distribution (i.e., asymptotic normality) holds

$$p_{\text{obs}}^{1/2}\left(\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}_0\right) \xrightarrow{d} \mathcal{N}\left(0, \boldsymbol{\zeta}\right)$$

In addition, we can justify Algorithm 1 in the main manuscript from a pure Bayesian point of view. Let $\boldsymbol{r}$ be an indicator matrix of whether elements in $\boldsymbol{Y}$ are observed or missing. Based on $\boldsymbol{r}$, we can define two subset vectors in $\boldsymbol{Y}$: $\boldsymbol{Y}_{\boldsymbol{r}}$ and $\boldsymbol{Y}_{\bar{\boldsymbol{r}}}$, which includes the observed and missing elements of $\boldsymbol{Y}$ respectively under the given missing structure $\boldsymbol{r}$. Now, given a prior $p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ for the multivariate normal mean and variance, the posterior distribution can be computed by

$$p\left(\boldsymbol{\mu}, \boldsymbol{\Sigma} \mid \mathbf{r}, \mathbf{Y}_{\mathbf{r}}\right) = \frac{p(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \int_{\mathcal{Y}_{\bar{r}}} p(\mathbf{r} \mid \mathbf{Y}, \phi) p(\mathbf{Y} \mid \theta) d\mathbf{Y}_{\bar{\mathbf{r}}}}{p\left(\mathbf{r}, \mathbf{Y}_{\mathbf{r}}\right)}. \tag{.1}$$

Further evaluating (.1) will give

$$\begin{aligned}
p\left(\boldsymbol{\mu}, \boldsymbol{\Sigma} \mid \mathbf{r}, \mathbf{Y}_{\mathbf{r}}\right) &= \int_{\mathcal{Y}_{\bar{\mathbf{r}}}} \frac{p(\boldsymbol{\mu}, \boldsymbol{\Sigma}) p(\mathbf{r} \mid \mathbf{Y}, \phi) p(\mathbf{Y} \mid \theta)}{p\left(\mathbf{r}, \mathbf{Y}_{\mathbf{r}}\right)} d\mathbf{Y}_{\bar{\mathbf{r}}} && (.2) \\
&= \int_{\mathcal{Y}_{\bar{r}}} \frac{p(\boldsymbol{\mu}, \boldsymbol{\Sigma}) p(\mathbf{r} \mid \mathbf{Y}, \phi) p(\mathbf{Y} \mid \theta)}{p\left(\mathbf{r}, \mathbf{Y}_{\mathbf{r}}\right)} \frac{p(\mathbf{r}, \mathbf{Y})}{p(\mathbf{r}, \mathbf{Y})} d\mathbf{Y}_{\bar{r}} && (.3) \\
&= \int_{\mathcal{Y}_{\bar{r}}} \frac{p(\boldsymbol{\mu}, \boldsymbol{\Sigma}) p(\mathbf{r} \mid \mathbf{Y}, \phi) p(\mathbf{Y} \mid \theta)}{p(\mathbf{r}, \mathbf{Y})} \frac{p(\mathbf{r}, \mathbf{Y})}{p\left(\mathbf{r}, \mathbf{Y}_{\mathbf{r}}\right)} d\mathbf{Y}_{\bar{\mathbf{r}}} && (.4) \\
&= \int_{\mathcal{Y}_{\bar{\mathbf{r}}}} p\left(\boldsymbol{\mu}, \boldsymbol{\Sigma} \mid \mathbf{r}, \mathbf{Y}_{\mathbf{r}}, \mathbf{Y}_{\bar{\mathbf{r}}}\right) p\left(\mathbf{Y}_{\bar{\mathbf{r}}} \mid \mathbf{r}, \mathbf{Y}_{\mathbf{r}}\right) d\mathbf{Y}_{\bar{\mathbf{r}}} && (.5) \\
&\approx \frac{1}{M} \sum_{m=1}^{M} p\left(\boldsymbol{\mu}, \boldsymbol{\Sigma} \mid \mathbf{r}, \mathbf{Y}_{\mathbf{r}}, \mathbf{Y}_{\bar{\mathbf{r}}}^{(m)}\right), \text{ with } \left\{\mathbf{Y}_{\bar{\mathbf{r}}}^{(m)}\right\}_m \overset{iid}{\sim} p\left(\mathbf{Y}_{\bar{\mathbf{r}}} \mid \mathbf{r}, \mathbf{Y}_{\mathbf{r}}\right) && (.6)
\end{aligned}$$

As we can see, Algorithm 1 is a size M Monte Carlo simulation from a purely Bayesian version and performs well with a proper choice of arguments.

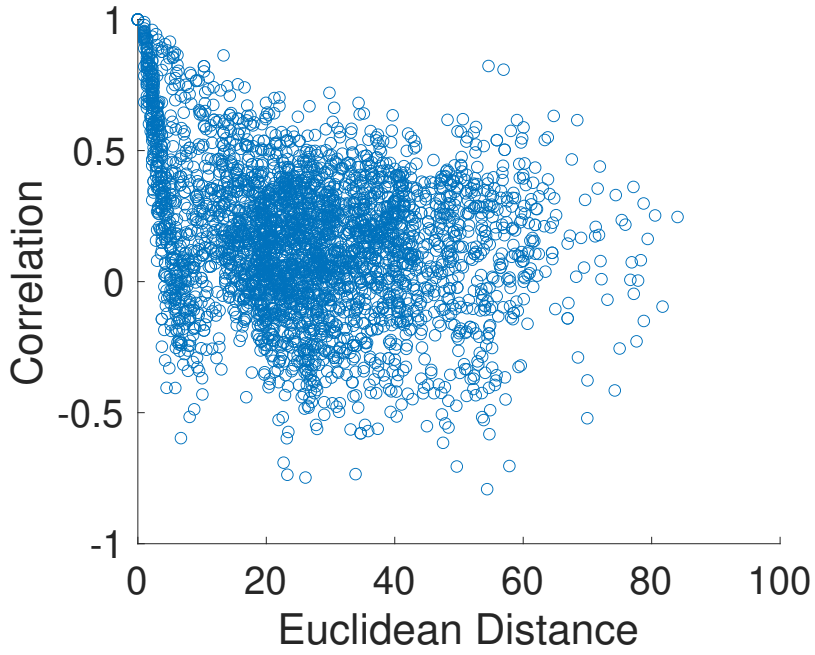


Figure B.1. *A scatter plot of voxel-pair correlations against voxel-pair spatial distance from a randomly selected region in a Kw imaging data.*

3C. Impropriety of NNGP in neuroimaging data imputation

We have discussed in the manuscript that nearest neighbor gaussian processes (NNGP) based imputation models are not applicable to brain imaging data. To visually demonstrate the reason for its inapplicability, we obtain the neurovascular water exchange (Kw) imaging data from a randomly selected brain region and plot the voxel-pair Kw correlations against voxel-pair spatial distance (Euclidean distance in 3D space). We see that the plotted points are approximately circular rather than linear, indicating that there is no direct relationship between voxel-pair correlation and spatial proximity, and thus NNGP-based imputation models are not appropriate for our application.

Appendix : MOAT

4A. Estimation of λ_1, λ_2

We implement the objective function (5) to extract multilayer sub-networks $\{B_c\}$ by Algorithm 1 with given λ_1 and λ_2 in the main manuscript. We can objectively select the optimal tuning parameters based on the likelihood function of B and G, where $B = (S, F; H)$ and $G = (V; F)$. Without loss of generality, we assume the weight $w(h_{(ij),k})$ in B is observed from a mixed Bernoulli distribution:

$$w(h_{(ij),k}) = I(\beta_{(ij),k} \neq 0) \sim \begin{cases} \text{Bernoulli}(\pi_1) & i,j,k \in B_c \\ \text{Bernoulli}(\pi_0) & \text{otherwise} \end{cases} \quad \text{with } \pi_1 > \pi_0. \tag{.1}$$

Given the reference distribution for $w(h_{(ij),k})$, we can obtain the likelihood function of λ_1, λ_2 since π_1, π_0 depends on λ_1, λ_2. In the case of π_1, the likelihood function can be written as

$$l_\lambda\left(\pi_1^{(\lambda_1,\lambda_2)} \mid w(h_{(ij),k})\right) = \sum_{(i,j,k)\in B_c} w(h_{(ij),k}) \log \pi_1^{(\lambda_1,\lambda_2)} \quad + \left(1 - w(h_{(ij),k})\right) \log\left(1 - \pi_1^{(\lambda_1,\lambda_2)}\right). \tag{.2}$$

The likelihood function of λ_1, λ_2 under π_0 can be constructed similarly. Next, λ_1, λ_2 can be

objectively determined by the Kullback–Leibler divergence with the reference mixed Bernoulli distribution based on the random graphs B and G:

$$
\begin{aligned}
&D_{KL}(P_{\cup_{c=1}^{\hat{C}} \hat{B}_c^{(\lambda_1,\lambda_2)}} \| Q) \\
&\quad = \left\{ \left[\sum_{i,j,k \in \hat{B}_c^{(\lambda_1,\lambda_2)}} \left(w(h_{(ij),k}) \pi_1 \log \frac{\pi_1}{\pi} + (1 - w(h_{(ij),k}))(1-\pi_1) \log \frac{(1-\pi_1)}{(1-\pi)} \right) \right.\right. \\
&\qquad \left.\left. + \sum_{i,j,k \notin \hat{B}_c^{(\lambda_1,\lambda_2)}} \left(w(h_{(ij),k}) \pi_0 \log \frac{\pi_0}{\pi} + (1 - w(h_{(ij),k}))(1-\pi_0) \log \frac{(1-\pi_0)}{(1-\pi)} \right) \right] \right\}, \quad (.3)
\end{aligned}
$$

where

$$
\pi := \frac{\sum_{i,j,k} I\left(a_{ijk} > \hat{r}\right)}{mn(n-1)/2}, \pi_1 := \frac{\sum_{(i,j,k) \in \hat{B}_c} I\left(a_{ijk} > \hat{r}\right)}{\left\| \hat{B}_c \right\|_0}, \pi_0 := \frac{\sum_{(i,j) \notin \hat{B}_c} I\left(a_{ijk} > \hat{r}\right)}{mn(n-1)/2 - \left\| \hat{B}_c \right\|_0},
$$

and $\left\| \hat{B}_c \right\|_0$ denotes the number of edges in the sub-graph $\hat{B}_c$. We select λ_1, λ_2 that most deviates from the null (i.e., B and G are random graphs).

4B. Proofs

4B.1. Proof of Lemma 1

First, suppose we observe a subgraph B_c from a random single-layer bipartite graph $B^*(S, F; H)$ with connection density p_1. We assume further that B_c is any subgraph with edge density $\frac{\sum_{ijk \in B^*} w(h_{(ij),k})}{|S_c||F_c|} \geq \gamma_1 \in (p_1, 1)$. Let $a_0, b_0 = \Omega\left(\max\left\{a^\epsilon, b^\epsilon\right\}\right)$ for some $0 < \epsilon < 1$. Then by Lemma 1 proposed in Wu et al. (2021), for sufficiently large a, b with $\zeta(\gamma_1, p_1) a_0 \geq 8 \log b$ and

$\zeta(\gamma_1, p_1)b_0 \geq 8\log a$, we have that

$$\mathbb{P}\left(|S_c| \geq a_0, |F_c| \geq b_0\right) \leq 2ab \cdot \exp\left(-\frac{1}{4}\zeta\left(\gamma_1, p_1\right) a_0 b_0\right),$$

where $\zeta(x, y) = \left\{\frac{1}{(x-y)^2} + \frac{1}{3(x-y)}\right\}^{-1}$.

Next, suppose we observe another subgraph G_c from a random single-layer unipartite graph $G^*(V; F)$ with connection density p_2. By Lemma 2.1 established in Chen et al. (2023), for any $\gamma_2 \in (p_2, 1)$, $v_0 = \omega(\sqrt{n})$, where ω denotes a looser lower bound, and n is large enough such that $\zeta(\gamma_2, p_2)v_0 \geq 4\log n$, we have

$$\mathbb{P}\left(|V_c| \geq n_0\right) \leq 2n \cdot \exp\left(-\frac{1}{4}\zeta\left(\gamma_2, p_2\right) v_0^2\right),$$

Now, we consider a subgraph B_c that is observed from a multilayer graph B with the first level B^* and the second level G^* as defined above. Suppose B_c also satisfies that (i) the edge density in B^* with $\frac{\sum_{ijk \in B_c} w(h_{(ij),k})}{|S_c||F_c|} \geq \gamma_1 \in (p_1, 1)$; (ii) the edge density in inner layer G with $\frac{\sum_{ij \in G_c} w(f_{ij})}{|V_c|^2} \geq \gamma_2 \in (p_2, 1)$. Then for sufficiently large m, n with $\zeta(\gamma_1, p_1)m_0 \geq 4\log n(n-1)$, $\zeta(\gamma_1, p_1)n_0(n_0-1) \geq 16\log m$, and $\zeta(\gamma_2, p_2)n_0 \geq 4\log n$, we have the following

$$\begin{aligned}
&\mathbb{P}\left(|S_c| \geq m_0, |F_c| \geq \frac{n_0(n_0-1)}{2}, |V_c| \geq n_0\right) \\
&\leq \mathbb{P}\left(|S_c| \geq m_0, |F_c| \geq \frac{n_0(n_0-1)}{2}\right) \cdot \mathbb{P}\left(|V_c| \geq n_0\right) \\
&\leq mn(n-1) \cdot \exp\left(-\frac{1}{4}\zeta\left(\gamma_1, p_1\right) m_0 n_0(n_0-1)/2\right) \cdot 2n \cdot \exp\left(-\frac{1}{4}\zeta\left(\gamma_2, p_2\right) n_0^2\right) \\
&\leq 2mn^2(n-1) \cdot \exp\left(-\frac{1}{8}\zeta\left(\gamma_1, p_1\right) m_0 n_0(n_0-1) - \frac{1}{4}\zeta\left(\gamma_2, p_2\right) n_0^2\right),
\end{aligned}$$

where $m_0, n_0 = \Omega\left(\max\left\{m^\epsilon, n^\epsilon\right\}\right)$ for some $0 < \epsilon < 1$.

4B.2. Proof of Theorem 1

To present the theoretical results, we consider the case where the weight $w(h_{(ij),k})$ is assigned to each SI-FC edge $h_{(ij),k}$ using the t statistic for testing whether $\beta_{(ij),k} = 0$ from the regression model (1). Other weight metrics (e.g., effect size, p-value) can be applied using the same proof technique shown below.

Let $\{h^1_{(ij),k}\}$ and $\{h^0_{(ij),k}\}$ be the sets of positive (e.g., SI-related) and negative edges, respectively, where $i \in [n] := (1, \ldots, n), j \in [n], k \in [m]$. We assume that weights $w(h^1_{(ij),k}) \overset{\text{iid}}{\sim} f_1$ and $w(h^0_{(ij),k}) \overset{\text{iid}}{\sim} f_0$, where f_1 and f_0 are two probability density functions with means and variances (p_1, σ_1^2) and (μ_0, σ_0^2), respectively. Since $w(h_{(ij),k})$ is determined by t statistic, we have $\mu_0 \leq p_1$ in general. Furthermore, let $w^1_{\text{min}} = \Omega(\{w(h^1_{(ij),k})\})$, where Ω denotes a loose lower bound; let $w^0_{\text{max}} = \varphi(\{w(h^0_{(ij),k})\})$, where φ denotes a loose upper bound. Lastly, consider a threshold $\theta \in (w^1_{\text{min}}, w^0_{\text{max}})$, we aim to prove Theorem 1 by investigating the asymptotic property of $\pi_0 := \mathbb{P}(w(h^0_{(ij),k}) > \theta)$ and $\pi_1 := \mathbb{P}(w(h^1_{(ij),k}) > \theta)$.

We focus on π_0 first. Let $Q_0 = \sum I(w(h^0_{(ij),k}) > \theta)$, $N_0 = |\{h^0_{(ij),k}\}|$ and $N_1 = |\{h^1_{(ij),k}\}|$.

For a fixed integer $x \in [\min(N_0, N_1)]$, we have

$$\begin{aligned}
\mathbb{P}(Q_0 \geq x) &= 1 - \mathbb{P}(Q_0 \leq x-1) \\
&= 1 - \sum_{i=1}^{x-1} \binom{N_0}{i} \pi_0^i (1-\pi_0)^{N_0 - i} \\
&= 1 - I_{1-\pi_0}(N_0 - x + 1, x) \\
&= 1 - \frac{\int_0^{1-\pi_0} t^{N_0 - x}(1-t)^{x-1} dt}{\int_0^1 t^{N_0 - x}(1-t)^{x-1} dt},
\end{aligned}$$

where $I(\cdot)$ is the regularized incomplete beta function. When the sample size $D \to \infty$, we have $\pi_0 \to \infty$, and thus

$$\mathbb{P}(Q_0 \geq x) \to 1 - 1 = 0. \tag{.4}$$

Similarly, let $Q_1 = \sum I(w(h^1_{(ij),k}) > \theta)$, we have

$$\begin{aligned}
\mathbb{P}(Q_1 < x) &= I_{1-\pi_1}(N_1 - x + 1, x) \\
&= \frac{\int_0^{1-\pi_1} t^{N_1 - x}(1-t)^{x-1} dt}{\int_0^1 t^{N_1 - x}(1-t)^{x-1} dt}.
\end{aligned}$$

When $D \to \infty$, we have $\pi_1 \to 1$, and thus

$$\mathbb{P}(Q_1 \geq x) \to 0. \tag{.5}$$

Let $\deg(k)$ denote the node degree of the k-th SI measure, i.e., $\deg(k) = \sum_{i,j} w(h_{(ij),k})$. Then

based on (.4) and (.5), we have

$$\mathbb{P}\left(\max\left(\deg(k^0)\right) > \min\left(\deg(k^1)\right)\right) \to 0, \tag{.6}$$

where k^0 and k^1 respectively represent negative and positive node of SI measure (e.g., node that has a significant structural impact on FC). Next, let $\mathcal{M}^*_{\text{row}}$ be the true membership of SI nodes (i.e., whether a given SI node is in $\{k^0\}$ or $\{k^1\}$). Let $\hat{\mathcal{M}}_{\text{row}}$ be the SI node membership estimated by optimizing the objective function in the main text using Algorithm 1. Based on (.6) and the property of the proposed greedy peeling algorithm which iteratively excludes nodes with the currently smallest degree, we have

$$\mathbb{P}(|\mathcal{M}^*_{\text{row}} - \hat{\mathcal{M}}_{\text{row}}| < \epsilon) = 1, \tag{.7}$$

where ϵ is an arbitrarily small number. Following the same logic, we can also obtain

$$\mathbb{P}(|\mathcal{M}^*_{\text{col}} - \hat{\mathcal{M}}_{\text{col}}| < \epsilon) = 1, \tag{.8}$$

where $\mathcal{M}^*_{\text{col}}$ and $\hat{\mathcal{M}}_{\text{col}}$ denote the true and estimated membership of FC nodes, respectively. Lastly, let $\mathcal{M}^*$ and $\hat{\mathcal{M}}$ be the true and estimated edge memberships (i.e., the community index of edges falling in the dense sub-networks). Since $\mathcal{M}^* = \mathcal{M}^*_{\text{row}} \bigotimes \mathcal{M}^*_{\text{col}}$ and $\hat{\mathcal{M}} = \hat{\mathcal{M}}_{\text{row}} \bigotimes \hat{\mathcal{M}}_{\text{col}}$, where $\bigotimes$ denotes the edge connecting two node sets on its left and right side. (.7) and (.8) jointly leads to

$$\mathbb{P}(|\mathcal{M}^* - \hat{\mathcal{M}}| < \epsilon) = 1, \tag{.9}$$

when the sample size $D \to \infty$.

4C. Additional information on real imaging data

4C.1. UK Biobank imaging data collection and preprocessing

The UK Biobank imaging data were collected using three sites each equipped with a Siemens Skyra 3T scanner and 32-channel RF head coil with high resolution T1-weighted (resolution = $1\times1\times1$ mm, FOV = $208\times256\times256$, duration = 5 min, 3D MPRAGE, sagittal, in-plane acceleration iPAT = 2, prescan-normalize). Diffusion data was acquired with the following parameters: a resolution = $2 \times 2 \times 2$ mm and two diffusion-weighted shells with all 100 distinct diffusion-encoding directions, $5b = 0$ images, $50xb = 1000$ and $2000s/mm^2$, FOV = $104 \times 104 \times 72$, and a 7-minute duration.

In order to ensure the quality of the fMRI data, we followed a series of outlined in Section 3.11, Page 18 of the Brain Imaging Documentation. Initially, we obtained the pre-processed rs-fMRI data for each subject from the file named "filtered_func_data_clean" provided by UK Biobank. The data underwent several processing steps: 1. Motion correction was performed using MCFLIRT (Jenkinson et al., 2002) to address any motion artifacts; 2. The entire 4D dataset was normalized for grand-mean intensity using a single multiplicative factor; 3. Highpass temporal filtering was applied using Gaussian-weighted least-squares straight line fitting with $\sigma = 50.0$s; 4. EPI unwarping was conducted using field maps as described earlier; 5. GDC unwarping was performed; 6. Finally, structured artifacts were eliminated using Independent Component Analysis (ICA) followed by FMRIB's ICA-based X-noisier (Beckmann and Smith, 2004; Salimi-Khorshidi et al., 2014). After completing these steps, the pre-processed data was obtained with

controlled quality (Alfaro-Almagro et al., 2018): Leave-one-out testing demonstrated that non-artifact components had a mean/median classification accuracy of 99.1/100.0%, while artifact components had a classification accuracy of 98.1/98.3%. We next applied a transformation to map functional data into a standard space (MNI 152) and applied a brain mask to the transformed data. Finally, we extracted times series from each voxel of interest in the masked data to calculate voxel-wise functional connectivity.

4C.2. Imaging data confounder control

The UK Biobank uses consistent scanner hardware and software across all its sites, eliminating the need for any adjustments to the scanner or acquisition protocol parameters. In our study, we considered a conventional set of confounding factors commonly used in typical imaging research. These confounds have been utilized and explained in the UK Biobank and Enigma Projects (Miller et al., 2016) and have been recognized as "conventional simple" confounds in Alfaro-Almagro et al. (2021). The confounders include age, sex, head size, site, and head motion. Furthermore, we incorporated educational level and Body Mass Index (BMI) as they are known to be associated with the strength of brain functional connectivity (Arenaza-Urquijo et al., 2013; Kullmann et al., 2012). Notably, since head size and motion estimates were already corrected during the pre-processing steps, we did not consider them as covariates in the regression model (1) in the main text.

C.3. Identified FAs

In this UK Biobank neuroimnaging data analysis, we collected 39 white matter FAs obtained from DTI processed using ENIGMA DTI protocols (Jahanshad et al., 2013) and labeled based on

the JHU ICBM DTI-81 Atlas (Mori et al., 2008; Smith et al., 2006). This atlas originally contains 48 regions, and 9 regions were removed following the ENIGMA protocols, resulting in the final 39 regions. Table 1 shows the complete list of the 39 regional white matter tracts.

Table. A list of 39 regional white matter (WM) integrity measured by fractional antitropy (FA).

Region of Interest		Hemisphere
Inferior cerebellar peduncle	ICP	/
Genu of corpus callosum	GCC	/
Body of corpus callosum	BCC	/
Splenium of corpus callosum	SCC	/
Fornix	FX	/
Corticospinal tract	CST-L	Left
Corticospinal tract	CST-R	Right
Anterior limb of internal capsule	ALIC-L	Left
Anterior limb of internal capsule	ALIC-R	Right
Posterior limb of internal capsule	PLIC-L	Left
Posterior limb of internal capsule	PLIC-R	Right
Retrolenticular part of internal capsule	RLIC-L	Left
Retrolenticular part of internal capsule	RLIC-R	Right
Anterior corona radiata	ACR-L	Left
Anterior corona radiata	ACR-R	Right
Superior corona radiata	SCR-L	Left
Superior corona radiata	SCR-R	Right
Posterior corona radiata	PCR-L	Left
Posterior corona radiata	PCR-R	Right
Posterior thalamic radiation	PTR-L	Left
Posterior thalamic radiation	PTR-R	Right
Sagittal stratum	SS-L	Left
Sagittal stratum	SS-R	Right
External capsule	EX-L	Left
External capsule	EX-R	Right
Cingulum cingulate gyrus	CGC-L	Left
Cingulum cingulate gyrus	CGC-R	Right
Cingulum hippocampus	CHG-L	Left
Cingulum hippocampus	CHG-R	Right
Fornix (cres)/stria terminalis	FXST-L	Left
Fornix (cres)/stria terminalis	FXST-R	Right
Superior longitudinal fasciculus	SLF-L	Left
Superior longitudinal fasciculus	SLF-R	Right
Superior fronto-occipital fasciculus	SFO-L	Left
Superior fronto-occipital fasciculus	SFO-R	Right
Uncinate fasciculus	UN-L	Left
Uncinate fasciculus	UN-R	Right
Tapetum	TAP-L	Left
Tapetum	TAP-R	Right

Table C.1: *White matter tracts defined following the ENIGMA protocols*

Bibliography

Agosta, F., Sala, S., Valsasina, P., Meani, A., Canu, E., Magnani, G., Cappa, S. F., Scola, E., Quatto, P., Horsfield, M. A., et al. (2013). Brain network connectivity assessed using graph theory in frontotemporal dementia. *Neurology*, 81(2):134–143.

Agustí, A., García-Pardo, M. P., López-Almela, I., Campillo, I., Maes, M., Romaní-Pérez, M., and Sanz, Y. (2018). Interplay between the gut-brain axis, obesity and cognitive function. *Frontiers in neuroscience*, 12:155.

Alfaro-Almagro, F., Jenkinson, M., Bangerter, N. K., Andersson, J. L., Griffanti, L., Douaud, G., Sotiropoulos, S. N., Jbabdi, S., Hernandez-Fernandez, M., Vallee, E., et al. (2018). Image processing and quality control for the first 10,000 brain imaging datasets from uk biobank. *Neuroimage*, 166:400–424.

Alfaro-Almagro, F., McCarthy, P., Afyouni, S., Andersson, J. L., Bastiani, M., Miller, K. L., Nichols, T. E., and Smith, S. M. (2021). Confound modelling in uk biobank brain imaging. *NeuroImage*, 224:117002.

Arenaza-Urquijo, E. M., Landeau, B., La Joie, R., Mevel, K., Mézenge, F., Perrotin, A., Desgranges, B., Bartrés-Faz, D., Eustache, F., and Chételat, G. (2013). Relationships between years of education and gray matter volume, metabolism and functional connectivity in healthy elders. *Neuroimage*, 83:450–457.

Arnold, S. and Trojanowski, J. (1996). Recent advances in defining the neuropathology of schizophrenia. *Acta neuropathologica*, 92(3):217–231.

Azur, M. J., Stuart, E. A., Frangakis, C., and Leaf, P. J. (2011). Multiple imputation by chained equations: what is it and how does it work? *International journal of methods in psychiatric research*, 20(1):40–49.

Bahrami, M., Laurienti, P. J., and Simpson, S. L. (2019). Analysis of brain subnetworks within the context of their whole-brain networks. *Human brain mapping*, 40(17):5123–5141.

Bai, F., Zhang, Z., Watson, D. R., Yu, H., Shi, Y., Yuan, Y., Qian, Y., and Jia, J. (2009). Abnormal integrity of association fiber tracts in amnestic mild cognitive impairment. *Journal of the neurological sciences*, 278(1-2):102–106.

Ball, G., Aljabar, P., Nongena, P., Kennea, N., Gonzalez-Cinca, N., Falconer, S., Chew, A. T., Harper, N., Wurie, J., Rutherford, M. A., et al. (2017). Multimodal image analysis of clinical influences on preterm brain development. *Annals of neurology*, 82(2):233–246.

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

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57(1):289–300.

Bischof, G. N. and Park, D. C. (2015). Obesity and aging: Consequences for cognition, brain structure and brain function. *Psychosomatic medicine*, 77(6):697.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008.

Bowman, F. D., Zhang, L., Derado, G., and Chen, S. (2012). Determining functional connectivity using fmri data with diffusion-based anatomical weighting. *NeuroImage*, 62(3):1769–1779.

Box, G. E. and Tiao, G. C. (2011). *Bayesian inference in statistical analysis*. John Wiley & Sons.

Brand, J. (1999). *Development, implementation and evaluation of multiple imputation strategies for the statistical analysis of incomplete data sets*.

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 theoretical analysis of structural and functional systems. *Nature reviews neuroscience*, 10(3):186–198.

Cano, C., Adarve, L., López, J., and Blanco, A. (2007). Possibilistic approach for biclustering microarray data. *Computers in biology and medicine*, 37(10):1426–1436.

Cao, M., Wang, J.-H., Dai, Z.-J., Cao, X.-Y., Jiang, L.-L., Fan, F.-M., Song, X.-W., Xia, M.-R., Shu, N., Dong, Q., et al. (2014). Topological organization of the human brain functional connectome across the lifespan. *Developmental cognitive neuroscience*, 7:76–93.

Cetin, M. S., Christensen, F., Abbott, C. C., Stephen, J. M., Mayer, A. R., Cañive, J. M., Bustillo, J. R., Pearlson, G. D., and Calhoun, V. D. (2014). Thalamus and posterior temporal lobe show greater inter-network connectivity at rest and across sensory paradigms in schizophrenia. *Neuroimage*, 97:117–126.

Chachlakis, D. G., Prater-Bennette, A., and Markopoulos, P. P. (2019). L1-norm tucker tensor decomposition. *IEEE Access*, 7:178454–178465.

Chaddock-Heyman, L., Erickson, K. I., Voss, M. W., Powers, J. P., Knecht, A. M., Pontifex, M. B., Drollette, E. S., Moore, R. D., Raine, L. B., Scudder, M. R., et al. (2013). White matter microstructure is associated with cognitive control in children. *Biological psychology*, 94(1):109–115.

Champion, C. J. (2003). Empirical bayesian estimation of normal variances and covariances. *Journal of multivariate analysis*, 87(1):60–79.

Chekuri, C., Quanrud, K., and Torres, M. R. (2022). Densest subgraph: Supermodularity, iterative peeling, and flow. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1531–1555. SIAM.

Chen, S., Bowman, F. D., and Mayberg, H. S. (2016). A bayesian hierarchical framework for modeling brain connectivity for neuroimaging data. *Biometrics*, 72(2):596–605.

Chen, S., Wu, Q., and Hong, L. E. (2019). Graph combinatorics based group-level network inference. *bioRxiv*, page 758490.

Chen, S., Zhang, Y., Wu, Q., Bi, C., Kochunov, P., and Hong, L. E. (2023). Identifying covariate-related subnetworks for whole-brain connectome analysis. *Biostatistics*, page kxad007.

Cheung, V., Cheung, C., McAlonan, G., Deng, Y., Wong, J., Yip, L., Tai, K., Khong, P., Sham, P., and Chua, S. (2008). A diffusion tensor imaging study of structural dysconnectivity in never-medicated, first-episode schizophrenia. *Psychological medicine*, 38(6):877–885.

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

Corrigan, F., Grand, D., and Raju, R. (2015). Brainspotting: Sustained attention, spinothalamic tracts, thalamocortical processing, and the healing of adaptive orientation truncated by traumatic experience. *Medical Hypotheses*, 84(4):384–394.

Craddock, R. C., James, G. A., Holtzheimer III, P. E., Hu, X. P., and Mayberg, H. S. (2012). A whole brain fmri atlas generated via spatially constrained spectral clustering. *Human brain mapping*, 33(8):1914–1928.

Cui, L.-B., Liu, J., Wang, L.-X., Li, C., Xi, Y.-B., Guo, F., Wang, H.-N., Zhang, L.-C., Liu, W.-M., He, H., et al. (2015). Anterior cingulate cortex-related connectivity in first-episode schizophrenia: a spectral dynamic causal modeling study with functional magnetic resonance imaging. *Frontiers in Human Neuroscience*, 9:589.

Derado, G., Bowman, F. D., and Kilts, C. D. (2010). Modeling the spatial and temporal dependence in fmri data. *Biometrics*, 66(3):949–957.

Dilsizian, S. E. and Siegel, E. L. (2014). Artificial intelligence in medicine and cardiac imaging: harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. *Current cardiology reports*, 16:1–8.

Doucet, A., Godsill, S. J., and Robert, C. P. (2002). Marginal maximum a posteriori estimation using markov chain monte carlo. *Statistics and Computing*, 12(1):77–84.

Drevets, W. C., Price, J. L., and Furey, M. L. (2008). Brain structural and functional abnormalities in mood disorders: implications for neurocircuitry models of depression. *Brain structure and function*, 213:93–118.

Eckelman, W. C., Reba, R. C., and Kelloff, G. J. (2008). Targeted imaging: an important biomarker for understanding disease progression in the era of personalized medicine. *Drug discovery today*, 13(17-18):748–759.

Efron, B. (2012). *Large-scale inference: empirical Bayes methods for estimation, testing, and prediction*, volume 1. Cambridge University Press.

Eickhoff, S. B., Thirion, B., Varoquaux, G., and Bzdok, D. (2015). Connectivity-based parcellation: Critique and implications. *Human brain mapping*, 36(12):4771–4792.

Enders, C. K. (2022). *Applied missing data analysis*. Guilford Publications.

Erdogmus, D. (2002). *Information theoretic learning: Renyi's entropy and its applications to adaptive system training*. University of Florida.

Ersche, K. D., Barnes, A., Jones, P. S., Morein-Zamir, S., Robbins, T. W., and Bullmore, E. T. (2011). Abnormal structure of frontostriatal brain systems is associated with aspects of impulsivity and compulsivity in cocaine dependence. *Brain*, 134(7):2013–2024.

Fan, J., Han, X., and Gu, W. (2012). Estimating false discovery proportion under arbitrary covariance dependence. *Journal of the American Statistical Association*, 107(499):1019–1035.

Fan, J. and Lv, J. (2008). Sure independence screening for ultrahigh dimensional feature space. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(5):849–911.

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., Laird, A. R., et al. (2016). The human brainnetome atlas: a new brain atlas based on connectional architecture. *Cerebral cortex*, 26(8):3508–3526.

Fedota, J. R. and Stein, E. A. (2015). Resting-state functional connectivity and nicotine addiction: prospects for biomarker development. *Annals of the new York Academy of Sciences*, 1349(1):64–82.

Ferri, J., Ford, J., Roach, B., Turner, J., Van Erp, T., Voyvodic, J., Preda, A., Belger, A., Bustillo, J., O'Leary, D., et al. (2018). Resting-state thalamic dysconnectivity in schizophrenia and relationships with symptoms. *Psychological medicine*, 48(15):2492–2499.

Fornito, A., Zalesky, A., and Bullmore, E. (2016). *Fundamentals of brain network analysis*. Academic Press.

Gaznick, N., Tranel, D., McNutt, A., and Bechara, A. (2014). Basal ganglia plus insula damage yields stronger disruption of smoking addiction than basal ganglia damage alone. *nicotine & tobacco research*, 16(4):445–453.

Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., and Rubin, D. (2014). Bayesian data analysis, vol. 2 crc press. *Boca Raton, FL*.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (1995). *Bayesian data analysis*. Chapman and Hall/CRC.

Geman, S. and Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, (6):721–741.

Grünwald, P. D. (2007). *The minimum description length principle*. MIT press.

Harrison, P. J. (2004). The hippocampus in schizophrenia: a review of the neuropathological evidence and its pathophysiological implications. *Psychopharmacology*, 174(1):151–162.

Hartigan, J. A. (1972). Direct clustering of a data matrix. *Journal of the american statistical association*, 67(337):123–129.

Hayden, E. P., Wiegand, R. E., Meyer, E. T., Bauer, L. O., O'connor, S. J., Nurnberger Jr, J. I., Chorlian, D. B., Porjesz, B., and Begleiter, H. (2006). Patterns of regional brain activity in alcohol-dependent subjects. *Alcoholism: Clinical and Experimental Research*, 30(12):1986–1991.

Hochreiter, S., Bodenhofer, U., Heusel, M., Mayr, A., Mitterecker, A., Kasim, A., Khamiakova, T., Van Sanden, S., Lin, D., Talloen, W., et al. (2010). Fabia: factor analysis for bicluster acquisition. *Bioinformatics*, 26(12):1520–1527.

Honey, C. J., Thivierge, J.-P., and Sporns, O. (2010). Can structure predict function in the human brain? *Neuroimage*, 52(3):766–776.

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of educational psychology*, 24(6):417.

J. K. Gupta, S. Singh, N. K. V. (2013). Mtba: Matlab toolbox for biclustering analysis. pages 94–97. IEEE.

Jahanshad, N., Kochunov, P. V., Sprooten, E., Mandl, R. C., Nichols, T. E., Almasy, L., Blangero, J., Brouwer, R. M., Curran, J. E., de Zubicaray, G. I., et al. (2013). Multi-site genetic analysis of diffusion images and voxelwise heritability analysis: A pilot project of the enigma–dti working group. *Neuroimage*, 81:455–469.

Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*, 17(2):825–841.

Jolliffe, I. T. and Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences*, 374(2065):20150202.

Kamvar, S., Klein, D., and Manning, C. (2003). Spectral learning. Technical Report 2003-25, Stanford InfoLab.

Kemmer, P. B., Wang, Y., Bowman, F. D., Mayberg, H., and Guo, Y. (2018). Evaluating the strength of structural connectivity underlying brain functional networks. *Brain Connectivity*, 8(10):579–594.

Kong, D., An, B., Zhang, J., and Zhu, H. (2019). L2rm: Low-rank linear regression models for high-dimensional matrix responses. *Journal of the American Statistical Association*.

Kullmann, S., Heni, M., Veit, R., Ketterer, C., Schick, F., Häring, H.-U., Fritsche, A., and Preissl, H. (2012). The obese brain: association of body mass index and insulin sensitivity with resting state network functional connectivity. *Human brain mapping*, 33(5):1052–1061.

Lambin, P., Leijenaar, R. T., Deist, T. M., Peerlings, J., De Jong, E. E., Van Timmeren, J., Sanduleanu, S., Larue, R. T., Even, A. J., Jochems, A., et al. (2017). Radiomics: the bridge between medical imaging and personalized medicine. *Nature reviews Clinical oncology*, 14(12):749–762.

Li, K.-H. (1988). Imputation using markov chains. *Journal of Statistical Computation and Simulation*, 30(1):57–79.

Li, R., Zhong, W., and Zhu, L. (2012a). Feature screening via distance correlation learning. *Journal of the American Statistical Association*, 107(499):1129–1139.

Li, Y., Long, J., He, L., Lu, H., Gu, Z., and Sun, P. (2012b). A sparse representation-based algorithm for pattern localization in brain imaging data analysis. *PloS one*, 7(12):e50332.

Lin, D., Zhang, J., Li, J., Calhoun, V. D., Deng, H.-W., and Wang, Y.-P. (2013). Group sparse canonical correlation analysis for genomic data integration. *BMC bioinformatics*, 14(1):1–16.

Loewe, K., Grueschow, M., Stoppel, C. M., Kruse, R., and Borgelt, C. (2014). Fast construction of voxel-level functional connectivity graphs. *BMC neuroscience*, 15(1):1–13.

Lu, T., Zhang, Y., Kochunov, P., Hong, E., and Chen, S. (2023). Network method for voxel-pair-level brain connectivity analysis under spatial-contiguity constraints. *arXiv preprint arXiv:2305.01596*.

Luxburg, U., Bousquet, O., and Belkin, M. (2004). Limits of spectral clustering. *Advances in neural information processing systems*, 17.

Lynall, M.-E., Bassett, D. S., Kerwin, R., McKenna, P. J., Kitzbichler, M., Muller, U., and Bullmore, E. (2010). Functional connectivity and brain networks in schizophrenia. *Journal of Neuroscience*, 30(28):9477–9487.

Marek, S., Tervo-Clemmens, B., Calabro, F. J., Montez, D. F., Kay, B. P., Hatoum, A. S., Donohue, M. R., Foran, W., Miller, R. L., Hendrickson, T. J., et al. (2022). Reproducible brain-wide association studies require thousands of individuals. *Nature*, 603(7902):654–660.

Margulis, L. and Sagan, D. (2000). *What is life?* Univ of California Press.

Mbatchou, J., Barnard, L., Backman, J., Marcketta, A., Kosmicki, J. A., Ziyatdinov, A., Benner, C., O'Dushlaine, C., Barber, M., Boutkov, B., et al. (2021). Computationally efficient whole-genome regression for quantitative and binary traits. *Nature genetics*, 53(7):1097–1103.

McClernon, F. J., Conklin, C. A., Kozink, R. V., Adcock, R. A., Sweitzer, M. M., Addicott, M. A., Chou, Y.-h., Chen, N.-k., Hallyburton, M. B., and DeVito, A. M. (2016). Hippocampal and insular response to smoking-related environments: neuroimaging evidence for drug-context effects in nicotine dependence. *Neuropsychopharmacology*, 41(3):877–885.

Mihalik, A., Chapman, J., Adams, R. A., Winter, N. R., Ferreira, F. S., Shawe-Taylor, J., Mourão-Miranda, J., Initiative, A. D. N., et al. (2022). Canonical correlation analysis and partial least squares for identifying brain-behaviour associations: a tutorial and a comparative study. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*.

Milchenko, M. and Marcus, D. (2013). Obscuring surface anatomy in volumetric imaging data. *Neuroinformatics*, 11:65–75.

Miller, K. L., Alfaro-Almagro, F., Bangerter, N. K., Thomas, D. L., Yacoub, E., Xu, J., Bartsch, A. J., Jbabdi, S., Sotiropoulos, S. N., Andersson, J. L., et al. (2016). Multimodal population brain imaging in the uk biobank prospective epidemiological study. *Nature neuroscience*, 19(11):1523–1536.

Moran, L. V., Sampath, H., Kochunov, P., and Hong, L. E. (2013a). Brain circuits that link schizophrenia to high risk of cigarette smoking. *Schizophrenia bulletin*, 39(6):1373–1381.

Moran, L. V., Tagamets, M. A., Sampath, H., O'Donnell, A., Stein, E. A., Kochunov, P., and Hong, L. E. (2013b). Disruption of anterior insula modulation of large-scale brain networks in schizophrenia. *Biological psychiatry*, 74(6):467–474.

Mori, S., Oishi, K., Jiang, H., Jiang, L., Li, X., Akhter, K., Hua, K., Faria, A. V., Mahmood, A., Woods, R., et al. (2008). Stereotaxic white matter atlas based on diffusion tensor imaging in an icbm template. *Neuroimage*, 40(2):570–582.

Mueller, S., Wang, D., Pan, R., Holt, D. J., and Liu, H. (2015). Abnormalities in hemispheric specialization of caudate nucleus connectivity in schizophrenia. *JAMA psychiatry*, 72(6):552–560.

Mulugeta, G., Eckert, M. A., Vaden, K. I., Johnson, T. D., and Lawson, A. B. (2017). Methods for the analysis of missing data in fmri studies. *Journal of biometrics & biostatistics*, 8(1).

Nichols, T. E. (2012). Multiple testing corrections, nonparametric methods, and random field theory. *Neuroimage*, 62(2):811–815.

Nichols, T. E. and Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human brain mapping*, 15(1):1–25.

Palaniyappan, L., White, T. P., and Liddle, P. F. (2012). The concept of salience network dysfunction in schizophrenia: from neuroimaging observations to therapeutic opportunities. *Current topics in medicinal chemistry*, 12(21):2324–2338.

Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., and Nichols, T. E. (2011). *Statistical parametric mapping: the analysis of functional brain images*. Elsevier.

Poldrack, R. A. (2008). The role of fmri in cognitive neuroscience: where do we stand? *Current opinion in neurobiology*, 18(2):223–227.

Pradat, P.-F. and Dib, M. (2009). Biomarkers in amyotrophic lateral sclerosis: facts and future horizons. *Molecular diagnosis & therapy*, 13:115–125.

Reiss, A. L., Mazzocco, M. M., Greenlaw, R., Freund, L. S., and Ross, J. L. (1995). Neurodevelopmental effects of x monosomy: a volumetric imaging study. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, 38(5):731–738.

Rogers, B. P., Morgan, V. L., Newton, A. T., and Gore, J. C. (2007). Assessing functional connectivity in the human brain by fmri. *Magnetic resonance imaging*, 25(10):1347–1357.

Royston, P. and White, I. R. (2011). Multiple imputation by chained equations (mice): implementation in stata. *Journal of statistical software*, 45:1–20.

Rubin, D. B. (2004). *Multiple imputation for nonresponse in surveys*, volume 81. John Wiley & Sons.

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

Salimi-Khorshidi, G., Douaud, G., Beckmann, C. F., Glasser, M. F., Griffanti, L., and Smith, S. M. (2014). Automatic denoising of functional mri data: combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage*, 90:449–468.

Sarpal, D. K., Robinson, D. G., Lencz, T., Argyelan, M., Ikuta, T., Karlsgodt, K., Gallego, J. A., Kane, J. M., Szeszko, P. R., and Malhotra, A. K. (2015). Antipsychotic treatment and functional connectivity of the striatum in first-episode schizophrenia. *JAMA psychiatry*, 72(1):5–13.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. CRC press.

Schafer, J. L. and Yucel, R. M. (2002). Computational strategies for multivariate linear mixed-effects models with missing values. *Journal of computational and Graphical Statistics*, 11(2):437–457.

Schaid, D. J., Chen, W., and Larson, N. B. (2018). From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nature Reviews Genetics*, 19(8):491–504.

Shabalin, A. A., Weigman, V. J., Perou, C. M., and Nobel, A. B. (2009). Finding large average submatrices in high dimensional data. *The Annals of Applied Statistics*, pages 985–1012.

Sheffield, J. M. and Barch, D. M. (2016). Cognition and resting-state functional connectivity in schizophrenia. *Neuroscience & Biobehavioral Reviews*, 61:108–120.

Simpson, S. L., Bowman, F. D., and Laurienti, P. J. (2013). Analyzing complex functional brain networks: Fusing statistics and network science to understand the brain. *Statistics Surveys*, 7:1–36.

Siuly, S. and Zhang, Y. (2016). Medical big data: neurological diseases diagnosis through medical data analysis. *Data Science and Engineering*, 1:54–64.

Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., Watkins, K. E., Ciccarelli, O., Cader, M. Z., Matthews, P. M., et al. (2006). Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage*, 31(4):1487–1505.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., et al. (2004). Advances in functional and structural mr image analysis and implementation as fsl. *Neuroimage*, 23:S208–S219.

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., et al. (2015). Uk biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS medicine*, 12(3):e1001779.

Sutherland, M. T. and Stein, E. A. (2018). Functional neurocircuits and neuroimaging biomarkers of tobacco use disorder. *Trends in molecular medicine*, 24(2):129–143.

Tae, W.-S., Ham, B.-J., Pyun, S.-B., Kang, S.-H., and Kim, B.-J. (2018). Current clinical applications of diffusion-tensor imaging in neurological disorders. *Journal of Clinical Neurology*, 14(2):129–140.

Tanner, M. A. and Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American statistical Association*, 82(398):528–540.

Thirion, B., Flandin, G., Pinel, P., Roche, A., Ciuciu, P., and Poline, J.-B. (2006). Dealing with the shortcomings of spatial normalization: Multi-subject parcellation of fmri datasets. *Human brain mapping*, 27(8):678–693.

Tsiatis, A. A. (2006). Multiple imputation: A frequentist perspective. *Semiparametric theory and missing data*, pages 339–374.

Tustison, N. J., Cook, P. A., Klein, A., Song, G., Das, S. R., Duda, J. T., Kandel, B. M., van Strien, N., Stone, J. R., Gee, J. C., et al. (2014). Large-scale evaluation of ants and freesurfer cortical thickness measurements. *Neuroimage*, 99:166–179.

Uurtio, V., Bhadra, S., and Rousu, J. (2019). Large-scale sparse kernel canonical correlation analysis. In *International Conference on Machine Learning*, pages 6383–6391. PMLR.

Vaden Jr, K. I., Gebregziabher, M., Kuchinsky, S. E., and Eckert, M. A. (2012). Multiple imputation of missing fmri data in whole brain analysis. *Neuroimage*, 60(3):1843–1855.

Verellen, D., De Ridder, M., Tournel, K., Duchateau, M., Reynders, T., Gevaert, T., Linthout, N., and Storme, G. (2008). An overview of volumetric imaging technologies and their quality assurance for igrt. *Acta Oncologica*, 47(7):1271–1278.

Vounou, M., Nichols, T. E., Montana, G., Initiative, A. D. N., et al. (2010). Discovering genetic associations with high-dimensional neuroimaging phenotypes: A sparse reduced-rank regression approach. *Neuroimage*, 53(3):1147–1159.

Wang, D., Zhou, Y., Zhuo, C., Qin, W., Zhu, J., Liu, H., Xu, L., and Yu, C. (2015). Altered functional connectivity of the cingulate subregions in schizophrenia. *Translational psychiatry*, 5(6):e575–e575.

Wang, H., Nie, F., Huang, H., Kim, S., Nho, K., Risacher, S. L., Saykin, A. J., Shen, L., and Initiative, A. D. N. (2012). Identifying quantitative trait loci via group-sparse multitask regression and feature selection: an imaging genetics study of the adni cohort. *Bioinformatics*, 28(2):229–237.

Wang, H., Nie, F., Huang, H., Risacher, S., Ding, C., Saykin, A. J., and Shen, L. (2011). Sparse multi-task regression and feature selection to identify brain imaging predictors for memory performance. In *2011 International Conference on Computer Vision*, pages 557–562. IEEE.

Wang, Z., Novikov, A., Zolna, K., Merel, J. S., Springenberg, J. T., Reed, S. E., Shahriari, B., Siegel, N., Gulcehre, C., Heess, N., et al. (2020). Critic regularized regression. *Advances in Neural Information Processing Systems*, 33:7768–7778.

Wehrle, F. M., Lustenberger, C., Buchmann, A., Latal, B., Hagmann, C. F., O'Gorman, R. L., and Huber, R. (2020). Multimodal assessment shows misalignment of structural and functional thalamocortical connectivity in children and adolescents born very preterm. *Neuroimage*, 215:116779.

Wetherill, R. R., Jagannathan, K., Shin, J., and Franklin, T. R. (2014). Sex differences in resting state neural networks of nicotine-dependent cigarette smokers. *Addictive behaviors*, 39(4):789–792.

Wieling, M. and Nerbonne, J. (2009). Bipartite spectral graph partitioning to co-cluster varieties and sound correspondences in dialectology. In *Proceedings of the 2009 Workshop on Graph-based Methods for Natural Language Processing (TextGraphs-4)*, pages 14–22.

Wig, G. S., Laumann, T. O., and Petersen, S. E. (2014). An approach for parcellating human cortical areas using resting-state correlations. *Neuroimage*, 93:276–291.

Witten, D. M., Tibshirani, R., and Hastie, T. (2009). A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis. *Biostatistics*, 10(3):515–534.

Woo, C.-W., Krishnan, A., and Wager, T. D. (2014). Cluster-extent based thresholding in fmri analyses: pitfalls and recommendations. *Neuroimage*, 91:412–419.

Wu, G.-R., Stramaglia, S., Chen, H., Liao, W., and Marinazzo, D. (2013). Mapping the voxel-wise effective connectome in resting state fmri. *PloS one*, 8(9):e73670.

Wu, Q., Huang, X., Culbreth, A. J., Waltz, J. A., Hong, L. E., and Chen, S. (2022). Extracting brain disease-related connectome subgraphs by adaptive dense subgraph discovery. *Biometrics*, 78(4):1566–1578.

Wu, Q., Zhang, Y., Huang, X., Ma, T., Hong, L. E., Kochunov, P., and Chen, S. (2021). A multivariate to multivariate approach for voxel-wise genome-wide association analysis. *bioRxiv*.

Wu, T., Wang, L., Hallett, M., Chen, Y., Li, K., and Chan, P. (2011). Effective connectivity of brain networks during self-initiated movement in parkinson's disease. *Neuroimage*, 55(1):204–215.

Wylie, K. P. and Tregellas, J. R. (2010). The role of the insula in schizophrenia. *Schizophrenia research*, 123(2-3):93–104.

Xia, Y. and Li, L. (2017). Hypothesis testing of matrix graph model with application to brain connectivity analysis. *Biometrics*, 73(3):780–791.

Zalesky, A., Fornito, A., and Bullmore, E. T. (2010). Network-based statistic: identifying differences in brain networks. *Neuroimage*, 53(4):1197–1207.

Zhang, J., Sun, W. W., and Li, L. (2023). Generalized connectivity matrix response regression with applications in brain connectivity studies. *Journal of Computational and Graphical Statistics*, 32(1):252–262.

Zhang, J., Wang, H., Zhao, Y., Guo, L., and Du, L. (2022). Identification of multimodal brain imaging association via a parameter decomposition based sparse multi-view canonical correlation analysis method. *BMC bioinformatics*, 23(3):1–14.

Zhou, H. and Li, L. (2014). Regularized matrix regression. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, 76(2):463.

Zhou, H., Li, L., and Zhu, H. (2013). Tensor regression with applications in neuroimaging data analysis. *Journal of the American Statistical Association*, 108(502):540–552.

Zhu, H., Khondker, Z., Lu, Z., and Ibrahim, J. G. (2014). Bayesian generalized low rank regression models for neuroimaging phenotypes and genetic markers. *Journal of the American Statistical Association*, 109(507):977–990.

Zhu, X., Suk, H.-I., Wang, L., Lee, S.-W., Shen, D., Initiative, A. D. N., et al. (2017). A novel relational regularization feature selection method for joint regression and classification in ad diagnosis. *Medical image analysis*, 38:205–214.

Zhuang, X., Yang, Z., Curran, T., Byrd, R., Nandy, R., and Cordes, D. (2017). A family of locally constrained cca models for detecting activation patterns in fmri. *NeuroImage*, 149:63–84.