Statistical Methods for Brain Connectivity and Dynamic fMRI Analysis

by

Junting Wang

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Statistics)
in The University of Michigan
2026

Doctoral Committee:

Associate Professor Kean Ming Tan, Co-Chair
Professor Ji Zhu, Co-Chair
Professor Jian Kang
Professor Gongjun Xu

Junting Wang
juntingw@umich.edu
ORCID iD: 0009-0003-2042-4708

© Junting Wang 2026

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisors, Ji Zhu and Kean Ming Tan, for their invaluable guidance, patience, and constant support throughout my doctoral studies. Their mentorship has shaped my understanding of research and has helped me grow both academically and personally. I am sincerely grateful for the time and effort they have invested in my development. I feel very fortunate to have had mentors who guided me through this journey and helped me reach this point. I would also like to extend my sincere thanks to my committee members, Gongjun Xu and Jian Kang, for their insightful comments, constructive feedback, and encouragement during the evaluation of this work. Their perspectives have helped strengthen this dissertation in important ways. I am also thankful to the faculty and staff in the Department of Statistics for their patience, professionalism, and support throughout my studies. Finally, I would like to thank my family and friends for their companionship, understanding, and encouragement. Their support gave me the courage and perseverance to continue, especially during challenging times. Completing this Ph.D. dissertation has been a long, difficult, yet deeply meaningful journey. I am profoundly grateful to the many individuals who have walked alongside me and made this achievement possible.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ii

LIST OF FIGURES v

LIST OF TABLES ix

LIST OF APPENDICES x

ABSTRACT xi

CHAPTER

1 Introduction **1**

2 Higher-Order Statistics for Constructing Centered Edge Functional Connectivity **5**

2.1 Introduction 5
2.2 Centered Edge Functional Connectivity (ceFC) 7
2.2.1 Sample Moment Estimator of ceFC 9
2.2.2 Multiple Hypothesis Testing 10
2.2.3 Thresholded Estimator of ceFC 13
2.2.4 Connection with nFC 15
2.2.5 Comparison with eFC 16
2.3 Numerical Studies 17
2.3.1 Estimation 17
2.3.2 Comparison with eFC 20
2.3.3 Hypotheses Testing on ceFC 21
2.4 An Application to the Midnight Scan Club Data 24
2.4.1 Identifying Overlapping Communities 25
2.4.2 Stability of the Proposed Estimators 31

3 Assessing the Predictive Value of Centered Edge Functional Connectivity for Phenotypic Traits **34**

3.1 Introduction 34

3.2 Methods . . . 36
3.2.1 AOMIC ID1000 Data . . . 36
3.2.2 Node-centric Functional Connectivity (nFC) . . . 37
3.2.3 Centered Edge Functional Connectivity (ceFC) . . . 38
3.2.4 Dimension Reduction . . . 38
3.2.5 Model . . . 40
3.3 Results . . . 42
3.3.1 Linear Regression Results . . . 43
3.3.2 PLS Results . . . 47
3.3.3 Explanation . . . 51

4 Spatio-temporally Varying Coefficient Model via Deep Neural Networks with Applications in fMRI . . . 55

4.1 Introduction . . . 55
4.2 Model and Estimation . . . 58
4.2.1 Image Response Regression with Spatio-Temporally Varying Coefficients . . . 58
4.2.2 Varying Coefficients Approximation . . . 59
4.2.3 Estimation Procedure . . . 63
4.3 Numerical Studies . . . 65
4.3.1 Spatial Design . . . 66
4.3.2 Temporal Design . . . 68
4.3.3 Spatio-Temporal Design . . . 68
4.4 Data Application . . . 71

5 Conclusion . . . 78

APPENDICES . . . 80

BIBLIOGRAPHY . . . 104

LIST OF FIGURES

FIGURE

2.1 Relative Frobenius norm estimation error for the non-sparse (top) and sparse (bottom) ceFC matrix with the number of brain parcels set to $p = 50$ (left) or $p = 200$ (right). The colored lines correspond to the proposed sample moment estimator (—), the thresholded estimators using the tuning parameter $C = 0.002$ (--), $C = 0.005$ (--), and $C = 0.01$ (--). The sample size n (number of time frames), number of brain parcels p, and tuning parameter C determine the thresholding τ through $\tau = C\sqrt{\log^4(TN^4)/T}$ for the thresholded estimator. 19

2.2 The true and false positive rates for sparse ceFC matrix with different choices of tuning parameter $C = 0.0001$ to $C = 1$. The colored curves correspond to the proposed thresholded estimator under different sample sizes n. 20

2.3 Relative Frobenius norm estimation error for a non-sparse ceFC matrix. The colored lines correspond to the proposed sample moment estimation for ceFC (—) and eFC (--). 22

2.4 Brain-region projection summaries for different ceFC estimators 29

2.5 Proportions of edges within each community and associated brain systems as depicted in Figure 2.4. The plot highlights the top three brain systems that demonstrate a significantly greater proportion of edges in pairwise comparisons with other systems. $***$ denotes the system with the highest count of significant outcomes, $**$ represents the second highest, and $*$ signifies the third highest. 30

2.6 Brain systems with significant proportions from each average estimator under the significance level of 0.05. The black box represents the insignificant result, and the yellow box represents the significant result. 31

2.7 The false positive and false negative rates are averaged across 16 systems (first row) or 5 communities (second row). The solid points on line graphs represent the average results over subjects, and each small transparent point corresponds to each subject's result. 33

3.1 Sum of test MSEs using 5 nFC PCs and varying numbers of ceFC PCs. The gray bars depict standard errors across 100 random splits. 44
3.2 Test MSE for each phenotypic trait. The green, orange, and blue bars represent results using 5 nFC PCs only, 5 nFC PCs with 1 ceFC PC, and training response mean only, respectively. The highlighted red bar denotes the case in which adding the ceFC PC yields the smallest test MSE. The highlighted purple bar denotes the case in which only using the training response mean is the best, even though adding ceFC PCs yields a smaller test error in comparison to using nFC PCs only. 45
3.3 Mean test Pearson correlation using 40 nFC PCs and varying numbers of ceFC PCs. The gray bars depict standard errors across 100 random splits. 46
3.4 Test Pearson correlation for each phenotypic trait. The green and orange bars represent results using 40 nFC PCs only, and 40 nFC PCs with 1 ceFC PC, respectively. The highlighted red bar denotes the case in which adding the ceFC PC yields a larger test correlation. 47
3.5 Sum of test MSEs over all phenotypic traits with standard errors in gray bars across 100 random splits based on 2 PLS components, 45 nFC PCs and varying numbers of ceFC PCs. 48
3.6 Test MSE for each phenotypic trait. The green, orange, and blue bars represent results using 45 nFC PCs only, 45 nFC PCs with 1 ceFC PC, and training response mean only, respectively. The highlighted red bar denotes the case in which adding the ceFC PC yields the smallest test MSE. The highlighted purple bar denotes the case in which only using the training response mean is the best, even though adding ceFC PC yields a smaller test error in comparison to using nFC PCs only. 49
3.7 Mean test Pearson correlation using 100 nFC PCs and varying numbers of ceFC PCs under 5 PLS components. The gray bars depict standard errors across 100 random splits. 50
3.8 Test Pearson correlation for each phenotypic trait with 5 PLS components. The green and orange bars represent results using 100 nFC PCs only, and 100 nFC PCs with 1 ceFC PC, respectively. The highlighted red bar denotes the case in which adding the ceFC PCs yields a larger test correlation. 51
3.9 R^2 from the linear regression of each true ceFC PC on all nFC PCs with 10 nodes. 53
3.10 R^2 from the regressing each ceFC PC, extracted from the ceFC estimate, on all nFC PCs using the full set of 200 brain nodes. 54

4.1 Relative MSE over time for the four spatio-temporal simulation settings: (a) additive, (b) multiplicative, (c) concatenative, and (d) concatenative with a moving activation region. In panel (a), the solid curves correspond to additive, concatenative, and transformer-based ST-IRRNN. In panel (b), they correspond to multiplicative, concatenative, and transformer-based ST-IRRNN. In panels (c) and (d), they correspond to concatenative and transformer-based ST-IRRNN. The dashed curves represent the fixed-window IRRNN with temporal window sizes $\{3, 5, 10\}$. 71
4.2 Relative test MSE over time for concatenative ST-IRRNN, transformer-based ST-IRRNN, and IRRNN with a time window size of 100 when predicting observed image responses from the estimated main effects in the Sherlock movie-watching data. 74
4.3 Covariate effect estimates from the concatenative ST-IRRNN (top row) and transformer-based ST-IRRNN (bottom row) for selected movie scenes. The three columns correspond to the speaking, number of people, and arousal covariates. Panels (a)–(f) display the Watson psychiatrist, Press conference, Poison vial #3, Watson and SH meet, Science of deduction conversation, and Sally did not make it home scenes, respectively. . 76

A.1 Counts of significant outcomes from pairwise comparisons of edge proportions among brain systems. The top three brain systems, identified by their top three highest counts within each method, are highlighted in Figure 4. 88
A.2 Normalized entropy of edge proportions in Figure 3 over all communities. The plot highlights the top three brain systems that demonstrate a significantly greater normalized entropy of edges in pairwise comparisons with other systems. $***$ denotes the system with the highest count of significant outcomes, $**$ represents the second highest, and $*$ signifies the third highest. 89
A.3 Counts of significant outcomes from pairwise comparisons of normalized entropy among brain systems. The top three brain systems, identified by their top three highest counts within each method, are highlighted in Figure A.2. 89
A.4 Sample moment estimation of eFC, ceFC, and eFC $-$ ceFC in (a), (b), and (c), respectively. Plot (d) shows the distribution of the overestimation ratio (%) on a log scale. 90

A.5 Correlation of each method from two halves of the simulation as split-half reliability. Plots (a) and (b) show the count of edge pairs for corresponding correlation values with $p = 50$ and $p = 200$ parcels, respectively. 91

B.1 Sum of validation MSEs over all phenotypic traits 94
B.2 Average validation Pearson correlations over all phenotypic traits with different numbers of nFC PCs and ceFC PCs. 94
B.3 Sum of validation MSEs over all phenotypic traits with different numbers of nFC PCs, ceFC PCs, and PLS components. 95
B.4 Average validation Pearson correlations over all phenotypic traits with different numbers of nFC PCs and ceFC PCs. 96
B.5 Sum of validation Spearman correlations over all phenotypic traits with different numbers of nFC PCs and ceFC PCs. 97
B.6 Mean test Spearman correlation using 55 nFC PCs and varying numbers of ceFC PCs. The gray bars depict standard errors across 100 random splits. 97
B.7 Test Spearman correlation for each phenotypic trait. The green and orange bars represent results using 55 FC PCs only and 55 FC PCs with 1 ceFC PC, respectively. The highlighted red bar denotes the case when adding the ceFC PC returns the largest test correlation. 98
B.8 Mean of validation Spearman correlations over all phenotypic traits with different numbers of nFC PCs and ceFC PCs. 99
B.9 Mean test Spearman correlation using 100 nFC PCs and varying numbers of ceFC PCs under 5 PLS components. The gray bars depict standard errors across 100 random splits. 99
B.10 Test Spearman correlation for each phenotypic trait with 5 PLS components. The green and orange bars represent results using 100 nFC PCs only and 100 nFC PCs with 1 ceFC PC, respectively. The highlighted red bar denotes the case when adding the ceFC PCd returns the largest test correlation. 100

C.1 A spatial design example for (a) a main effect $\beta_1(\boldsymbol{s}_v)$, (b) an individual effect $\alpha_1(\boldsymbol{s}_v)$, and (c) log variance of the error term $\log[\sigma^2(\boldsymbol{s}_v)]$. 102
C.2 Examples of spatio-temporal designs 102

LIST OF TABLES

TABLE

2.1 The type I error and power for testing the 1000 random off-diagonal entries in ceFC at the 0.05 significance level are averaged over 5000 data sets. The power study is separated into 1000 random weak signals $\in (0, 0.2]$ and 1000 random medium signals $\in (0.2, 0.5]$, respectively. Simulation results for $p = 50$ and $p = 200$ variables over a range of sample size n are reported. 23

2.2 The false discovery rate for testing 3000 random entries for $p = 50$ and 5000 random entries for $p = 200$ are averaged over 5000 data sets. These random subsets retain the sparsity in the corresponding ceFC matrix. The significance level α is set to be $\{0.05, 0.1, 0.2\}$ and the sample size n ranges from 400 to 700. 24

2.3 False positive rate (FPR) and false negative rate (FNR) are averaged across 16 brain systems and 5 communities over 10 subject data for each estimator. 32

LIST OF APPENDICES

APPENDIX

A Appendix of Chapter I **80**

B Appendix of Chapter II **92**

C Appendix of Chapter III **101**

ABSTRACT

Functional magnetic resonance imaging (fMRI) provides rich measurements of brain activity and has become an important tool for studying brain organization, individual variation, and brain-covariate associations. However, fMRI data are often high-dimensional, structurally complex, and heterogeneous across subjects, which poses challenges for statistical analysis. This dissertation develops flexible and interpretable statistical methodologies for analyzing complex fMRI data.

The first part of this work focuses on higher-order relationships in the brain. We introduce centered edge functional connectivity (ceFC) as an unbiased measure of covariation between brain edges, study its estimation under both classical and high-dimensional regimes, and develop a multiple hypothesis testing framework to identify nonzero edge-level associations while controlling the false discovery rate.

The second part investigates whether the proposed ceFC provides additional predictive information for phenotypic traits beyond conventional node-centric functional connectivity (nFC) within a brain-wide association study framework. The results show that when a sufficient set of information from nFC is included, features extracted from ceFC provide limited additional predictive benefit.

The third part studies image response regression under dynamic naturalistic stimuli. We propose a flexible spatio-temporal image response regression via neural networks (ST-IRRNN) to estimate dynamic coefficient functions. We also introduce several ST-IRRNN variants designed to accommodate different forms of spatio-temporal dependence. Numerical studies demonstrate improved estimation accuracy and temporal stability compared to fixed-window spatial baselines. Additionally, an application to a naturalistic movie-watching fMRI dataset shows that the proposed methods effectively recover scene-dependent effects of audio, visual, and emotion-related stimuli across large-scale brain systems. Overall, this dissertation contributes new statistical methods for studying both connectivity structure and dynamic effects in fMRI data. These tools provide new perspectives for understanding brain organizations and brain associations with phenotypes, behaviors, and continuous naturalistic stimuli.

CHAPTER 1

Introduction

Recent technological advancements, such as functional magnetic resonance imaging (fMRI), have significantly advanced our understanding of brain function through cost-efficient and non-invasive (albeit indirect) acquisition measures of brain activity (Wagner et al., 1998; Buckner et al., 1996). fMRI detects changes in spontaneous and task-free BOLD fluctuations under specific tasks or at rest, yielding an indirect measure of brain activity over time (Ogawa et al., 1990; Kwong et al., 1992). These measurements make it possible to investigate not only localized activation patterns, but also coordinated responses across distributed brain systems. As a result, fMRI has become a central tool in modern neuroscience and has contributed greatly to the study of brain organization, individual variation, cognitive function, and disease-related alterations in neural activity.

Neuroimaging measurements are often high-dimensional, highly structured, and heterogeneous across subjects. Moreover, the dependence structure in such data can be complicated, involving spatial organization, temporal evolution, network structure, and subject-level variation. These challenges motivate the development of methods that accommodate complex structure while remaining statistically principled and scientifically meaningful.

One major line of research in fMRI studies focuses on connectivity patterns among brain regions. In particular, functional connectivity (FC) is commonly defined through statistical dependence, often correlation, between BOLD signals from distinct brain regions (Friston et al., 1993; Aertsen et al., 1989; Friston, 1994). Numerous studies have applied FC in brain networks (Power et al., 2011a; Finn et al., 2015), shedding light on the brain's functional organization and identifying putative biomarkers for a range of diseases (Greicius et al., 2003; Wen et al., 2018; Yeo et al., 2011), including but not limited to autism spectrum disorders (Monk et al., 2009; Weng et al., 2010), Alzheimer's disease (Supekar et al., 2008), major depressive disorder (Yan et al., 2019; Kaiser et al., 2015), and mild cognitive impairment (Wang et al., 2020). In addition to such group-wise differences in functional connectivity

between healthy individuals and patients, the existing literature also reveals heterogeneous patterns in brain connectivity since each individual is unique (Mueller et al., 2013), holding the potential for personalized treatment (Finn and Constable, 2016; Finn et al., 2015).

Despite many successes, conventional node-centric FC, referred to here as nFC, is inherently limited to pairwise relationships among brain regions and may not fully describe more complicated structures. Recent edge-centric work suggests that higher-order connectivity can reveal overlapping community organization, individual differences, and clinically relevant network structure that are less visible under standard node-based summaries (Glickfeld and Olsen, 2017; Gatica et al., 2021; Luppi et al., 2022; Palla et al., 2005; Yeo et al., 2014; Beckmann et al., 2005; van den Heuvel and Sporns, 2013; Faskowitz et al., 2020; Esfahlani et al., 2020; Pope et al., 2021; Wen et al., 2022; Liu et al., 2022b; Chumin et al., 2022; Jo et al., 2021b,a; Ville et al., 2021; Dai et al., 2023; Yang et al., 2023; Zamani Esfahlani et al., 2022; Sun et al., 2023). Motivated by this perspective, Chapter 2 introduces centered edge functional connectivity (ceFC) as an unbiased measure of covariation between pairs of edges and studies statistical properties of its estimators. We propose a sample moment estimator in a classical regime with a sufficient sample size and a regularized estimator with a thresholding operator in a high-dimensional setting in which the number of variables is comparable to or exceeds the sample size. In addition, a multiple testing framework is developed to identify nonzero ceFC entries while controlling the false discovery rate. This chapter establishes statistical consistency properties of the proposed estimators and demonstrates their empirical performance through simulation studies. The application of proposed estimators to the Midnight Scan Club fMRI dataset further illustrates their capability to uncover overlapping community structures by edge clustering, providing a new perspective on brain organization.

In addition to network structure, another major objective within connectivity analysis is to understand how variation in brain connectivity relates to external covariates, such as demographic characteristics, behavioral traits, and phenotypic information across individuals (Zatorre et al., 2012; Rosenberg et al., 2018; Cheng et al., 2015; Krakauer et al., 2017; Ooi et al., 2022b).. This question has become a central component of brain-wide association studies (BWAS), where large-scale FC patterns are used to explain or predict a wide range of subject-level traits (Marek et al., 2022; Rosenberg and Finn, 2022).

Prior work on brain-phenotype association has often relied on conventional nFC measurements. On the other hand, the ceFC methodology developed in Chapter 2 has been shown to provide additional perspective on brain structure and organization. While this higher-order connectivity can capture interaction patterns that may not

be immediately visible from nFC, it remains unclear whether this new metric contains additional valuable information for phenotypic prediction.

Building on a BWAS framework (Marek et al., 2022), Chapter 3 examines whether ceFC provides additional predictive information for phenotypic traits beyond nFC. In particular, subject-level predictors are constructed from both conventional nFC and ceFC proposed in Chapter 2. That is, each node-pair or edge-pair connectivity is treated as a covariate, and regularization and dimension reduction are employed to address the high dimensionality within ceFC. The empirical results show that, when a sufficiently rich set of nFC features is included, adding ceFC features provides limited additional predictive benefit. Further analyses indicate that much of the dominant variation in ceFC can already be explained by conventional nFC components. These results refine our understanding of the role of higher-order connectivity in brain-phenotype prediction and suggest that increased representational complexity does not necessarily translate into improved predictive power.

Beyond studying connectivity patterns among brain regions in the first two parts of this dissertation, another central goal in fMRI studies is to characterize how brain imaging responses vary with external covariates, such as demographic characteristics, behavioral measurements, clinical scores, and stimulus-related features, across anatomical space. This problem is scientifically important because such associations can localize covariate effects over the brain image domain and may provide interpretable markers for brain function and disease (Zhu et al., 2014; Woo et al., 2017; Smith et al., 2009). Existing approaches include mass univariate methods, tensor-based models, and spatially varying coefficient models (SVCMs), each of which addresses different aspects of the complexity of image-valued responses (Friston, 2003; Sun and Li, 2017; Zhu et al., 2014; Li et al., 2017, 2021). More recently, Zhang et al. (2023) proposed image response regression via neural networks (IRRNN), which uses deep neural networks to approximate spatially varying coefficient functions while preserving coefficient-level interpretability.

Many contemporary imaging studies, however, arise in dynamic settings where the relevant covariate effects evolve over both space and time. This is especially important under naturalistic fMRI paradigms, such as movie watching, in which the stimuli of interest are continuous, multimodal, and context-dependent, closely resembling real-world conditions (Sonkusare et al., 2019; Finn and Bandettini, 2021; Matusz et al., 2019). In such settings, the effect of a stimulus-related covariate can vary over both space and time, and exhibit dependence patterns that are not adequately described by static models.

Chapter 4 addresses this setting by developing a spatio-temporally varying coefficient model (STVCM) for dynamic image-response data and proposing spatio-

temporal image response regression via neural networks (ST-IRRNN) for estimating the associated coefficient functions. Several ST-IRRNN variants are considered to accommodate different forms of spatio-temporal dependence, including additive, multiplicative, concatenative, and transformer-based formulations. This framework remains interpretable at the coefficient-function level while retaining the flexibility of modern deep learning methods. Simulation studies with different spatio-temporal structures show that the proposed methods achieve lower estimation error and greater temporal stability than fixed-window spatial baselines. A real-data application to the Sherlock movie-watching fMRI dataset (Chen et al., 2017) further illustrates that the fitted spatio-temporal coefficient maps reveal scene-dependent effects of auditory, visual-social, and emotion-related covariates across large-scale brain systems.

CHAPTER 2

Higher-Order Statistics for Constructing Centered Edge Functional Connectivity

2.1 Introduction

Conventional node-centric functional connectivity summarizes pairwise associations between brain regions. Although this representation has been highly useful, it may fail to capture more complex dependence structures among connections themselves. Previous studies suggest that there may be higher-order interactions involved in functional connectivity (Glickfeld and Olsen, 2017; Gatica et al., 2021; Luppi et al., 2022), which may be obscured when using field-standard node-based approaches. For instance, higher-order estimates of functional connectivity may provide a more detailed understanding of functional brain networks through pervasive, non-trivial overlapping community organizations (Palla et al., 2005; Yeo et al., 2014; Beckmann et al., 2005; van den Heuvel and Sporns, 2013). Various lines of evidence have indicated that brain regions participate in multiple functional roles (Hove and Martinez, 2024; Ackerman, 1992; Rishel et al., 2013), highlighting the importance of studying overlapping communities within brain networks, which can be trivially detected using an edge-centric approach within functional brain networks. In addition, higher-order connectivity in brain studies has been shown to provide valuable insights into various types of disorders, including but not limited to cognitive and neurological disorders, which may assist in the early detection and diagnosis of diseases (Plis et al., 2014; Chen et al., 2016; Zhang et al., 2017).

Recently, a novel concept of edge-centric functional connectivity (eFC), which examines connectivity between pairs of edges, was introduced in (Faskowitz et al., 2020). This edge-centric framework studies the co-fluctuations in the BOLD signals between pairs of edges and has been applied in brain networks, revealing a diverse community structure in which brain regions can be grouped into multiple communities based on edge clusters (Faskowitz et al., 2020). The identification of overlapping communities by edge connectivity provides a deeper understanding of brain organi-

zation (Chumin et al., 2022), where at least two edge communities can link all brain systems (Jo et al., 2021b). Recently, the edge-centric framework has been found to provide another perspective for understanding structural and functional features of brain networks (Esfahlani et al., 2020; Pope et al., 2021; Wen et al., 2022; Liu et al., 2022b), identifying individual differences (Jo et al., 2021a; Ville et al., 2021), serving as biomarkers for various disorders (Dai et al., 2023; Yang et al., 2023; Zamani Esfahlani et al., 2022; Sun et al., 2023), and uncovering the mechanism underlying therapeutic efficacy (Khambhati et al., 2021).

The proposed eFC measures the strength of connectivity between edge pairs based on the co-fluctuation of edge time series while also taking the absolute levels of edge time series, i.e., their mean values, into consideration. However, including the mean values of edge time series may bias the estimation of the association between edge pairs when the parameter of interest is their covariation. This can occur because the mean values can inflate or deflate the perceived strength of relationships between edges, leading to a skewed measurement and potentially obscuring the true underlying covariation between edges. To address this issue, the present study aims to provide an unbiased estimation of the covariation between edge pairs, denoted as ceFC, and to investigate the statistical properties of its estimators. Each brain parcel is treated as a variable in our analysis. For a single subject, one intuitive estimation of ceFC, denoted as $\mathbf{\Theta}$, can be achieved using the sample moment estimator, which consists of the empirical estimators of the fourth-order and second-order moments by definition. Under the assumption of sub-Gaussian variables, we study the rate of convergence of the sample moment estimator $\widehat{\mathbf{\Theta}}$ under the Frobenius norm. Additionally, we develop a multiple-hypothesis testing framework to examine the non-zero elements of the ceFC matrix while controlling the false discovery rate (FDR). Considering that different edge pairs can be dependent, our inference framework employs the Benjamini–Yekutieli procedure, a method designed to control FDR when accounting for dependencies between pairs of statistical tests (Benjamini and Yekutieli, 2001).

An important consideration for creating a reliable sample moment estimator is ensuring that the number of time frames is sufficient relative to the number of variables, i.e., the number of brain regions. In fMRI studies, there are multiple schemes to parcellate human brains into brain regions, or regions of interest (ROIs), where the common choices range from 50 up to 1000 parcels (Yeo et al., 2011; Glasser et al., 2016; Gordon et al., 2016; Power et al., 2011b; Schaefer et al., 2018). Different brain parcellations can impact functional connectivity magnitudes and generate differences among individuals (Bryce et al., 2021), while the number of time frames can be relatively limited in comparison to the high level of brain parcellations, particularly during specific tasks (Van Essen et al., 2013). It can be challenging when the number

of brain areas is comparable to or even greater than the number of time frames during scan sessions, as the sample estimator for covariance becomes singular, whereas the actual covariance matrix is always positive definite.

In high-dimensional settings, it has been shown that the sample covariance typically does not perform well (Marčenko and Pastur, 1967; Johnstone, 2001; Johnstone and Lu, 2009), in which the estimation error can fail to converge to zero. To address this issue, significant effort has been devoted to developing alternative estimators. There is a rich literature on estimating large covariance matrices through banding or tapering (Bickel and Levina, 2008b; Cai et al., 2010; Yi and Zou, 2013; Bien et al., 2016), thresholding (Bickel and Levina, 2008a; Karoui, 2008; Rothman et al., 2009; Cai and Liu, 2011), and penalization (Lam and Fan, 2009; Ravikumar et al., 2011). These techniques have been applied in various domains (Fan and Liu, 2013; Xu and Lindquist, 2015; Markon, 2010; Asai et al., 2020), including neuroscience. For example, people have employed thresholding operators to eliminate weak connections in brain networks and reveal distinct features of brain networks (Stam, 2004; De Vico Fallani et al., 2017; Garrison et al., 2015; Nicolini et al., 2020). In light of this, we adopt the thresholding operator $\mathbf{T}_\tau$ to truncate values in the sample moment estimator below a certain threshold τ to 0 as the thresholded estimator for a class of sparse ceFC matrices.

In this chapter, we aim to explore the unbiased estimators of covariation among edges under both general scenarios with sufficient sample sizes and the high-dimensional setting in which the sample size is smaller than the number of variables. We demonstrate that both types of estimators for centered eFC - the sample moment estimator and the proposed thresholded estimator - provide consistent results through statistical rates of convergence and simulation studies. Building on the sample moment estimator, we construct a multiple hypothesis testing framework to examine the sparsity of the ceFC matrix and demonstrate that this inference framework achieves asymptotic control of FDR. We also apply the proposed estimators to the Midnight Scan Club (MSC) dataset (Gordon et al., 2017) and investigate the underlying overlapping structures through edge clustering. By evaluating the clustering results from three types of estimators, we show that the FDR-controlled and thresholded estimators provide relatively more stable clustering outcomes compared to those based on the sample moment estimator.

2.2 Centered Edge Functional Connectivity (ceFC)

The covariance matrix has found wide applications across scientific disciplines including genomics (Pascal et al., 2008; Bulik-Sullivan et al., 2015; Duncan et al., 2017)

and neuroscience (Salvador et al., 2005; Schneidman et al., 2006; Chen et al., 2008). However, the covariance matrix is limited to modeling linear dependencies between pairs of variables.

In recent years, many authors in the neuroscience literature have found that computing higher-order relationships among variables can help facilitate a deeper understanding of the structural and functional brain organizations (Faskowitz et al., 2020; Esfahlani et al., 2020; Pope et al., 2021; Wen et al., 2022; Liu et al., 2022b), recognize individual differences (Jo et al., 2021a; Ville et al., 2021), and serve as biomarkers for various disorders (Dai et al., 2023; Yang et al., 2023; Zamani Esfahlani et al., 2022; Sun et al., 2023).

The primary goal of our study is to introduce and validate an unbiased framework for connectivity between edges. We begin by defining the key concepts underlying our analysis before presenting the main theoretical results.

At its core, traditional FC measures the statistical dependence between pairs of brain regions (e.g. correlation) (Friston et al., 1993; Aertsen et al., 1989; Friston, 1994). A more recent concept, eFC, extends this idea to a higher order by quantifying the co-fluctuation between pairs of edges themselves (Faskowitz et al., 2020). An edge time series can be constructed, for instance, by taking the element-wise product of the time series of two connected nodes. The original eFC is then calculated as a scaled measure of the co-fluctuation between two such edge time series (Faskowitz et al., 2020).

A critical limitation of the eFC formulation is its sensitivity to the mean of the edge time series (i.e., the static functional connectivity). This can introduce a sample-specific bias, potentially impacting the reproducibility of scientific findings. To address this, we propose ceFC, which is mathematically defined as the unbiased covariation between two edge time series after their means have been removed. By design, ceFC isolates the pure dynamic co-fluctuation between edges from the confounding influence of static connectivity.

Functional connectivity measures the magnitude of dependency between a pair of brain voxels or parcels. One common measurement is the Pearson correlation coefficient using their BOLD signal time series. Let $\boldsymbol{x}_j = [x_{1j}, \ldots, x_{nj}]$ denote the time series recorded from parcel j, where n is the number of time frames, and p is the number of parcels. For each parcel j, we compute the time-averaged mean $\mu_j = \frac{1}{n}\sum_i x_{ij}$ and standard deviation $\sigma_j = \sqrt{\frac{1}{n-1}\sum_i (x_{ij} - \mu_j)^2}$, such that we can obtain the z-scored time series vector $\boldsymbol{z}_j = \frac{x_j - \mu_j}{\sigma_j}$. For any pair of z-scored time series vectors $\boldsymbol{z}_j, \boldsymbol{z}_k$, the correlation between parcels j and k can be computed as $r_{jk} = \frac{1}{n-1}\sum_i [z_{ij} \cdot z_{ik}]$. By repeating this computation for all pairs of parcels, we obtain the node-by-node correlation matrix $\boldsymbol{R} = [r_{jk}]$ of size $p \times p$, which is the

estimate of node-centric functional connectivity (nFC).

Here, we treat brain parcels as random variables and use uppercase letters to denote them. Specifically, let $X_1, \ldots, X_p$ be the random variables for the p brain parcels, and $\boldsymbol{X} = (X_1, \ldots, X_p)^{\mathrm{T}}$ be an p-dimensional random vector with mean zero and an $p \times p$ covariance matrix $\boldsymbol{\Sigma} = [\sigma_{jk}] \in \mathbb{R}^{p \times p}$. The above computation for the correlation can be written as $r_{jk} = \frac{\mathbb{E}[X_j X_k]}{\sigma_j \sigma_k}$, where σ_j, σ_k are the standard deviations of X_j, X_k, respectively. We define ceFC between $\{X_j, X_k\}$ and $\{X_s, X_t\}$ as

$$\Theta_{jk,st} = \mathbb{E}[X_j X_k X_s X_t] - \mathbb{E}[X_j X_k] \cdot \mathbb{E}[X_s X_t]. \tag{2.2.1}$$

Here, $\Theta_{jk,st}$ measures linear dependency between two pairs of variables $\{X_j, X_k\}$ and $\{X_s, X_t\}$. If we consider the z-scored BOLD signals $Z_1, \ldots, Z_p$, $\Theta_{jk,st}$ can be viewed as the connection between edge jk and edge st by taking the product of centered edge time series $Z_j Z_k$ and $Z_s Z_t$. Moreover, if we standardize the edge time series $Z_j Z_k$ and $Z_s Z_t$ and then take the product, the resultant ceFC will be the correlation between edge jk and edge st and thus scaled between $[-1, 1]$. Considering all combinations of parcel pairs for entries in ceFC, Θ will be a $\frac{p(p-1)}{2} \times \frac{p(p-1)}{2}$ matrix.

In the following sections, we will propose two types of estimators for ceFC under low- and high-dimensional settings, respectively. We first examine a straightforward sample moment estimator, which is analogous to the original eFC but applied to mean-centered data. We will demonstrate its statistical limitations in high-dimensional settings, i.e., when the number of time frames is limited compared to the number of parcels. Next, we present a multiple-hypothesis testing framework to simultaneously evaluate the strength of each edge pair covariation, demonstrating that this inference procedure effectively controls the FDR. We then introduce and evaluate a more sophisticated thresholded estimator, designed to provide consistent and robust estimates even when the number of brain parcels is large relative to the number of time frames.

2.2.1 Sample Moment Estimator of ceFC

To estimate ceFC $\Theta_{jk,st}$ defined in (2.2.1), we consider the situation where we have a sufficient sample size n compared to the number of variables p. In other words, the number of time frames of BOLD signals is large enough compared to the number of brain parcels. Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ be n independently and identically distributed copies of $\boldsymbol{X}$, which represent the BOLD time series of whole brain parcels. A natural estimator of $\Theta_{jk,st}$ is constructed by the sample mean of each expectation term in (2.2.1), which

is the following sample moment-based estimator

$$\widehat{\Theta}_{jk,st} = \frac{1}{n}\sum_{i=1}^{n} X_{ij}X_{ik}X_{is}X_{it} - \frac{1}{n}\sum_{i=1}^{n} X_{ij}X_{ik} \cdot \frac{1}{n}\sum_{i=1}^{n} X_{is}X_{it}. \tag{2.2.2}$$

If we repeat this computation for all combinations of parcel pairs, we will obtain a $\frac{p(p-1)}{2} \times \frac{p(p-1)}{2}$ matrix $\widehat{\Theta}$, a sample estimate of Θ.

In the following, we establish non-asymptotic upper bounds for the estimation error of the sample moment estimator in (2.2.2) and show that the sample moment estimator converges to the true ceFC matrix within a bound as a function of $\sqrt{p^4 \log^4(np^4)/n}$, providing a consistent estimate when the sample size (or number of time frames) n is sufficiently large, specifically with n much larger than p^4. We start with imposing some conditions on the random vector $\boldsymbol{X}$.

Condition 2.1. The random vector $\boldsymbol{X} \in \mathbb{R}^p$ is *sub-Gaussian* with mean zero and covariance matrix $\boldsymbol{\Sigma} = \mathbb{E}[\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}] \in \mathbb{R}^{p\times p}$ that is positive definite. Consequently, for each X_j and $v \geq 1/4$, we have $\mathbb{E}[|X_j|^{4v}] \leq K^{4v}(4v)^{2v}$ for some constant $K > 0$.

Denote the Frobenius norm of a matrix $\mathbf{A}$ as

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i,j} a_{ij}^2}.$$

We consider the statistical rate of the sample moment estimator in (2.2.2) under the Frobenius norm.

Theorem 2.2.1. Under Condition 2.1, we have

$$\|\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}\|_F^2 \leq Cp^4 \frac{\log^4(np^4)}{n},$$

for some constants C, with probability at least $1 - n^{-1}$.

The proof can be found in Appendix A.1.4. We can see that with a fixed number of variables p, the estimation error measured under the Frobenius norm approaches zero with increasing sample size n. This convergent error demonstrates that our estimator is consistent and becomes increasingly accurate as more data are collected, ensuring reliable estimation of the true ceFC matrix in practical applications.

2.2.2 Multiple Hypothesis Testing

We now perform statistical inference on the ceFC matrix. Specifically, we are interested in simultaneously testing the hypothesis that each element in the ceFC matrix

is equal to zero. The literature acknowledges that not all brain regions or edges exhibit strong connections with others (Achard and Bullmore, 2007; Faskowitz et al., 2022; Bullmore and Sporns, 2012; Faskowitz et al., 2020), and such sparse network patterns can be a reflection of the trade-offs between the wiring costs and topological structures (Achard and Bullmore, 2007; Faskowitz et al., 2022; Bullmore and Sporns, 2012; Faskowitz et al., 2020). In light of these findings, our multiple hypothesis testing framework provides a statistical inference procedure to systematically evaluate the strength of each edge-pair connection within the ceFC matrix.

Formally, our goal is to test each l^{th} pair of $\{jk, st\}$ in $\mathbf{\Theta}$ simultaneously:

$$H_{0l} : \Theta_{jk,st} = 0 \ v.s. \ H_{1l} : \Theta_{jk,st} \neq 0, \quad l = 1, \ldots, N.$$

Given that the ceFC matrix $\mathbf{\Theta}$ is symmetric with diagonals equal to one, it is sufficient to examine the set of N upper-triangular entries of $\mathbf{\Theta}$ for the N hypotheses with $N = \frac{p(p-1)\ [p(p-1)-2\ ]}{8}$.

It remains to construct a test statistic for the multiple hypotheses. Based on the proposed sample moment estimator $\widehat{\Theta}_{jk,st}$, by an application of the multivariate central limit theorem and multivariate delta method, it can be shown that

$$\sqrt{n}\left[\widehat{\Theta}_{jk,st} - \Theta_{jk,st}\right] \xrightarrow{d} \mathcal{N}(0, \Lambda_{jk,st}), \tag{2.2.3}$$

where

$$\begin{aligned}\Lambda_{jk,st} = &\operatorname{Var}(X_jX_kX_sX_t) + \sigma_{st}^2 \operatorname{Var}(X_jX_k) + \sigma_{jk}^2 \operatorname{Var}(X_sX_t) \\ &- 2\sigma_{st}\operatorname{Cov}\left(X_jX_k, X_jX_kX_sX_t\right) \\ &- 2\sigma_{jk}\operatorname{Cov}(X_sX_t, X_jX_kX_sX_t) \\ &+ 2\sigma_{jk}\sigma_{st}\operatorname{Cov}(X_jX_k, X_sX_t).\end{aligned}$$

Under the null hypothesis, we define the test statistic as

$$\gamma_{jk,st} = \frac{\widehat{\Theta}_{jk,st}}{\sqrt{\widehat{\Lambda}_{jk,st}/n}}, \tag{2.2.4}$$

where $\widehat{\Lambda}_{jk,st}$ is the sample estimate of $\Lambda_{jk,st}$. For ease of notation, we denote $(X_{ij}X_{ik})_c = X_{ij}X_{ik} - n^{-1}\sum_{i=1}^n X_{ij}X_{ik}$, $(X_{is}X_{it})_c = X_{is}X_{it} - n^{-1}\sum_{i=1}^n X_{is}X_{it}$, and $(X_{ij}X_{ik}X_{is}X_{it})_c = X_{ij}X_{ik}X_{is}X_{it} - n^{-1}\sum_{i=1}^n X_{ij}X_{ik}X_{is}X_{it}$. We can write $\widehat{\Lambda}_{jk,st}$ as

$$\frac{1}{n-1}\sum_{i=1}^n (X_{ij}X_{ik}X_{is}X_{it})_c^2 + \left(\frac{1}{n}\sum_{i=1}^n X_{is}X_{it}\right)^2 \cdot \frac{1}{n-1}\sum_{i=1}^n (X_{ij}X_{ik})_c^2$$

$$
\begin{aligned}
&+ \left(\frac{1}{n}\sum_{i=1}^{n} X_{ij}X_{ik}\right)^2 \cdot \frac{1}{n-1}\sum_{i=1}^{n}(X_{is}X_{it})_c^2 \\
&- 2\left(\frac{1}{n}\sum_{i=1}^{n} X_{is}X_{it}\right) \cdot \frac{1}{n-1}\sum_{i=1}^{n}(X_{ij}X_{ik})_c(X_{ij}X_{ik}X_{is}X_{it})_c \\
&- 2\left(\frac{1}{n}\sum_{i=1}^{n} X_{ij}X_{ik}\right) \cdot \frac{1}{n-1}\sum_{i=1}^{n}(X_{is}X_{it})_c(X_{ij}X_{ik}X_{is}X_{it})_c \\
&+ 2\left(\frac{1}{n}\sum_{i=1}^{n} X_{ij}X_{ik}\right)\left(\frac{1}{n}\sum_{i=1}^{n} X_{is}X_{it}\right) \cdot \frac{1}{n-1}\sum_{i=1}^{n}(X_{ij}X_{ik})_c(X_{is}X_{it})_c.
\end{aligned}
$$

Let $\mathcal{R}_{\alpha_0 l}$ be the rejection region, then the corresponding p-value can be determined as

$$p_l = \inf_{\alpha_0} P_{0l}(\gamma_{jk,st} \in \mathcal{R}_{\alpha_0 l}).$$

Under the multiple testing setting, we now aim to control the false discovery rate (FDR)(Benjamini and Hochberg, 1995) defined as

$$\text{FDR} = \mathbb{E}\left[\frac{\#\text{ false positives}}{\#\text{ rejections}}\right],$$

under some predetermined level α. Given that the p-values p_l for different edge correlation entries can be dependent, we adopt the Benjamini–Yekutieli procedure (Benjamini and Yekutieli, 2001), which controls FDR among dependent hypotheses with a cut-off value for ordered p-values. Specifically, the FDR procedure can be summarized as Algorithm 1.

Algorithm 1 FDR Control Procedure (Benjamini and Yekutieli, 2001).

Input: A list of p-values $p_1 \cdots p_N$ for null hypotheses $H_{01}, \ldots, H_{0N}$ and a prespecified level α.

Initialization: Number of rejections $r = 0$.

1. Order p-values as $p_{(1)} \leq \cdots \leq p_{(N)}$ for corresponding $H_{0(1)}, \ldots, H_{0(N)}$.
2. Set $L = \sum_{l=1}^{N} 1/l$.
3. For $l = 1, \ldots, N$:

   If $p_{(l)} \leq \frac{l}{NL}\alpha$:

   Reject corresponding $H_{0(l)}$.

   Update $r = l$.

   Until $p_{(l)} > \frac{l}{NL}\alpha$.

Output: A set of rejected $H_{0(1)}, \ldots, H_{0(r)}$.

We demonstrate that the FDR can be controlled at a prespecified level α asymptotically by the above procedure in the following theorem. In other words, as the

number of time frames n increases, the FDR is asymptotically bounded above by some level α plus a term that converges to zero.

Theorem 2.2.2. For a pair of $\{jk, st\}$ and some level $\alpha \in (0, 1)$, the testing procedure above controls FDR at level α asymptotically that

$$\text{FDR} \leq \alpha + o(1).$$

The detailed proof can be found in Appendix A.1.6. Asymptotic control of the FDR implies that for large datasets, the proportion of false positives (i.e., incorrectly identified significant edge pairs) is maintained at or below the chosen significance level α, ensuring the reliability of the results in large-sample scenarios.

2.2.3 Thresholded Estimator of ceFC

In Section 2.2.1, we demonstrate that under the Frobenius norm, the upper bound of the sample moment estimation error behaves as a function of $p^2 \log^2(np^4)/\sqrt{n}$ and thus does not converge to zero under the high-dimensional setting with $p > n$, which can be a common situation in many domains. For example, the number of time frames n can be limited compared to the high level of brain parcellations p, especially under specific tasks (Van Essen et al., 2013). This high-dimensional scenario presents a fundamental statistical challenge: with limited time frames, many of the observed weak (low-weight) connections are statistically indistinguishable from random noise or sampling error. It becomes infeasible to reliably estimate the numerous parameters in the ceFC matrix without additional structural constraints. To address this challenge, we introduce a regularized estimator by truncating certain entry values below a specified threshold in the sample moment estimator to zero, thus providing a consistent estimation even when the number of parcels is greater than the number of time frames. This thresholding procedure is a necessary statistical step to mitigate the risk of overfitting and to ensure that subsequent analyses are based on reliable signals rather than spurious correlations. It has been a widely adopted strategy in the existing literature to achieve consistent estimation of large covariance matrices (Bickel and Levina, 2008a; Karoui, 2008; Rothman et al., 2009; Cai and Liu, 2011). Such a thresholding technique is often constructed based on structural assumptions that involve the sparsity of the target matrix and has been applied in various disciplines (Fan and Liu, 2013; Xu and Lindquist, 2015; Markon, 2010; Asai et al., 2020), including neuroscience (Honey et al., 2007; de Reus and van den Heuvel, 2013; Buchanan et al., 2020). For instance, the thresholding operator has been frequently employed in brain studies to eliminate spurious connections under the assumption that the likelihood of a connection being spurious decays monotonically with its weight, revealing

features of brain networks (Stam, 2004; De Vico Fallani et al., 2017; Garrison et al., 2015; Nicolini et al., 2020). To this end, we propose a regularized estimator through the thresholding operator, especially tailored for a class of sparse ceFC matrices. We first introduce the following notations.

For a matrix $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{p\times p}$, we define the thresholding operator T_τ as:

$$T_\tau(\mathbf{A}) = [a_{ij}\mathbb{1}(|a_{ij}| \geq \tau)].$$

With this operator $\mathbf{T}_\tau$, we truncate the entries a_{ij} in a matrix $\mathbf{A}$ to zero if their absolute values are less than a certain cut-off value τ.

We consider the class of sparse ceFC matrices that

$$\mathcal{U}_\tau(m) = \{\mathbf{\Theta} : \sum_{\{j,k,s,t\}} \mathbb{1}(\Theta_{jk,st} \neq 0) \leq m\}.$$

That is, the number of nonzero entries in such sparse ceFC matrices would be no more than certain values m. For such a sparse ceFC matrix $\mathbf{\Theta}$, we propose to estimate each entry $\Theta_{jk,st}$ by its thresholded estimator

$$T_\tau(\widehat{\Theta}_{jk,st}). \tag{2.2.5}$$

Similar to the sample moment estimation, which has been shown to provide a bounded statistical rate for a general edge correlation matrix, we study the bound of concentration inequality of the thresholded estimator with a certain choice of threshold τ. The proof can be found in Section A.1.5.

Theorem 2.2.3. Under Condition 2.1, uniformly on $\mathcal{U}_\tau(m)$, for some constant C, if $\tau = C\sqrt{\log^4(np^4)/n}$, we have

$$\|T_\tau(\widehat{\mathbf{\Theta}}) - \mathbf{\Theta}\|_F^2 \leq mC^2\frac{\log^4(np^4)}{n},$$

with probability at least $1 - n^{-1}$.

In situations where the true ceFC is sparse under the class $\mathcal{U}_\tau(m)$, the regularized estimator (2.2.5) approaches the true ceFC with an upper bound delineated by

$$\sqrt{\frac{m\log^4(np^4)}{n}},$$

which converges to zero even for $p > n$. On the other hand, by Theorem 2.2.1, the sample moment estimator (2.2.2) converges to the true edge correlation under the

bound as the function

$$\sqrt{\frac{p^4 \log^4(np^4)}{n}},$$

which fails to converge to zero when $p > n$. Accordingly, the proposed regularized estimator under the thresholding operator provides a consistent estimation for ceFC in the high-dimensional scenario. Moreover, even when the true ceFC is non-sparse, the regularized estimator can outperform the sample moment estimator with an appropriate tuning parameter when the sample size is small, which can be illustrated through the numerical studies in Section 2.3.1.

2.2.4 Connection with nFC

As the definition of ceFC (2.2.1) involves connectivity between nodes, it is of interest to study the connection between the ceFC matrix and the nFC matrix. Indeed, Magnus and Neudecker (1979) have shown the mathematical relationship between a covariance matrix and a ceFC-like matrix under a Gaussian distribution based on the commutation matrix. By definition, a commutation matrix $\boldsymbol{K}_{m,n}$ is a square $mn \times mn$ matrix consisting of mn submatrices and the (i^{th}, j^{th}) submatrix $\boldsymbol{K}_{i,j}$ contains 1 in its (j^{th}, i^{th}) entry and 0 otherwise. When $m = n$, we can simplify the notation $\boldsymbol{K}_{n,n}$ to $\boldsymbol{K}_n$. Specifically, Magnus and Neudecker (1979) state that for a Gaussian random vector $\boldsymbol{X} \in \mathbb{R}^N$ with mean $\widetilde{\boldsymbol{\mu}}$ and covariance $\widetilde{\boldsymbol{\Sigma}}$,

$$\mathrm{Var}(\boldsymbol{X} \otimes \boldsymbol{X}) = (\boldsymbol{I} + \boldsymbol{K}_p)(\widetilde{\boldsymbol{\Sigma}} \otimes \widetilde{\boldsymbol{\Sigma}} + \widetilde{\boldsymbol{\Sigma}} \otimes \widetilde{\boldsymbol{\mu}}\widetilde{\boldsymbol{\mu}}' + \widetilde{\boldsymbol{\mu}}\widetilde{\boldsymbol{\mu}}' \otimes \widetilde{\boldsymbol{\Sigma}}),$$

where $\otimes$ denotes the Kronecker product. Here $\mathrm{Var}(\boldsymbol{X} \otimes \boldsymbol{X})$ can be viewed as an extension of ceFC, which includes the covariance between the node pairs $\{X_j, X_k\}$ and $\{X_s, X_t\}$ and also the case when $j = k$ and/or $s = t$. Our study ignores the second case in $\boldsymbol{\Theta}$, as there is no edge connecting the node X_j itself. Based on the result in Magnus and Neudecker (1979), when the random vector $\boldsymbol{X} = (X_1, \ldots, X_p)^{\mathrm{T}}$ is Gaussian with mean $\boldsymbol{0}$ and a $p \times p$ covariance matrix $\boldsymbol{\Sigma} = [\sigma_{jk}] \in \mathbb{R}^{p \times p}$, we can rewrite

$$\boldsymbol{\Theta} = \ (\boldsymbol{I} + \boldsymbol{K}_{N,N-1}\ )\ (\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}_{-j,-k}\ ) \in \mathbb{R}^{p(p-1) \times p(p-1)} \tag{2.2.6}$$

for $j, k \in \{1, \ldots, p\}$. There are two modifications to Magnus and Neudecker (1979)'s result. First, $\boldsymbol{K}_{p,p-1}$ is a variation of the communication matrix $\boldsymbol{K}_p$, which contains $p \times (p-1)$ submatrices each with dimension of $(p-1) \times (p-1)$. In particular, the (j^{th}, k^{th}) submatrix contains 1 in its $(j-1^{th}, k^{th})$ entry when $j > k$ and $(j^{th}, (k-1)^{th})$ entry when $j < k$, and 0 elsewhere. Second, the matrix $\boldsymbol{\Sigma}_{-j,-k}$ removes σ_{jk} from $\boldsymbol{\Sigma}$. Specifically, $\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}_{-j,-k} = [\sigma_{jk}\boldsymbol{\Sigma}_{-j,-k}]$. Such modifications are mainly based on the removal of the term σ_{jj} for $j \in \{1, \ldots, p\}$, which does not retain practical meaning

in some application fields.

When considering the special case of z-scored node time series, equation (2.2.6) provides the mathematical relationship between ceFC and nFC, indicating that the ceFC matrix can be built from nFC. This may not be surprising as the construction of ceFC matrix depends on the edge time series, which in turn is obtained based on the node time series.

2.2.5 Comparison with eFC

The concept of ceFC is closely related to eFC as proposed by (Faskowitz et al., 2020). Both ceFC and eFC can quantify the connection between pairs of edges but from different perspectives. This distinction has profound implications for the interpretation and reproducibility of findings in edge-centric functional connectivity.

Specifically, eFC calculates the strength of edge connections based on the inner product of two edge time series, while ceFC is evaluated using centered edge time series. Mathematically, let E_{jk} denote the edge time series between parcels j and k, which can be obtained by taking the element-wise product of X_j and X_k as $(X_{1j}X_{1k}, \cdots, X_{nj}X_{nk})$. The root mean square of the edge time series can be defined as $\|E_{jk}\| = \sqrt{E_{jk}^T E_{jk}/n}$. The standard deviation of E_{jk} is $sd(E_{jk}) = \sqrt{(E_{jk} - \sigma_{jk})^T (E_{jk} - \sigma_{jk})/(n-1)}$, where σ_{jk} is the mean of the edge time series E_{jk}, or the covariance between parcels j and k with zero-mean node time series. By definition, eFC between edges jk and st can be computed as

$$eFC_{jk,st} = \frac{E_{jk}^T E_{st}/n}{\|E_{jk}\|\|E_{st}\|}.$$

Similarly, we can obtain ceFC between edges jk and st under the scale between -1 and 1 by

$$ceFC_{jk,st} = \frac{E_{jk}^T E_{st}/n - \sigma_{jk}\sigma_{st}}{sd(E_{jk})sd(E_{st})}.$$

The critical distinction lies in the centering step. The presence of the product of means, $\sigma_{jk}\sigma_{st}$, in the eFC formulation introduces a systematic bias that conflates the covariation between edge time series with the magnitudes of their individual means. Consequently, two edge pairs might exhibit high eFC not because their fluctuations are synchronized, but because both pairs have strong average connections. This confound makes eFC results dependent on sample-specific mean connectivity levels, threatening the reproducibility and generalization of findings across studies with different populations or acquisition protocols.

Therefore, even if eFC and ceFC were to yield visually similar structures in a given dataset as shown in 2.4.1, the communities identified by eFC are built upon a biased

foundation, where the similar patterns could be dependent on the specific datasets, while those from ceFC reflect a more reproducible signature of higher-order brain dynamics. In the special scenario where parcel pairs j, k and s, t are both uncorrelated ($\sigma_{jk} = 0$ and $\sigma_{st} = 0$) with sufficient time frames, the centering step becomes unnecessary and $|E_{jk}|$ equals $sd(E_{jk})$, making $eFC_{jk,st}$ equivalent to $ceFC_{jk,st}$ after standardization. However, this represents an edge case rather than the typical scenario in functional neuroimaging.

2.3 Numerical Studies

2.3.1 Estimation

In this section, we perform extensive numerical studies to assess the performance of the sample moment estimator and the regularized estimator discussed in Sections 2.2.1 and 2.2.3 under different scenarios. The proposed regularized estimator involves selecting a thresholding tuning parameter $\tau > 0$. Suggested by Theorem 2.2.3, we set $\tau = C \cdot \sqrt{\log^4(np^4)/n}$ with $C \in \{0.002, 0.005, 0.01\}$ for some weak to moderate controls under the simulation setting. To compare the performance across the two methods, we measure the estimation error between the proposed estimator $\widehat{\boldsymbol{\Theta}} = [\widehat{\Theta}_{jk,st}]$ and the true ceFC $\boldsymbol{\Theta}^* = [\Theta^*_{jk,st}]$ under the relative Frobenius norm $\|\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}^*\|_F / \|\boldsymbol{\Theta}^*\|_F = \sqrt{\sum_{jk,st}(\widehat{\Theta}_{jk,st} - \Theta^*_{jk,st})^2} / \sqrt{\sum_{jk,st}(\Theta^*_{jk,st})^2}$ for $\{jk, st\} \in \{1, \ldots, p(p-1)/2\}$. Moreover, we define the true positive rate (TPR) and false positive rate (FPR) as the proportion of correctly estimated non-zero entries and incorrectly estimated non-zero entries, respectively:

$$\mathrm{TPR}(\widehat{\boldsymbol{\Theta}}, \boldsymbol{\Theta}^*) = \frac{\#\{(jk, st) : \widehat{\Theta}_{jk,st} \neq 0 \text{ and } \Theta^*_{jk,st} \neq 0\}}{\#\{(jk, st) : \Theta^*_{jk,st} \neq 0\}}; \tag{2.3.1}$$

$$\mathrm{FPR}(\widehat{\boldsymbol{\Theta}}, \boldsymbol{\Theta}^*) = \frac{\#\{(jk, st) : \widehat{\Theta}_{jk,st} \neq 0 \text{ and } \Theta^*_{jk,st} = 0\}}{\#\{(jk, st) : \Theta^*_{jk,st} = 0\}}. \tag{2.3.2}$$

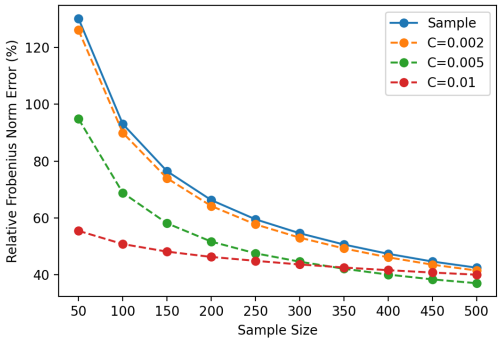
We start with generating n independent and identically distributed p-dimensional vectors from a multivariate normal distribution, i.e., $\boldsymbol{X}_i \sim N(\mathbf{0}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is some $p \times p$ covariance matrix that will be specified later in this section. Under the normality assumption, the true ceFC (2.2.1) is evaluated as $\Theta^*_{jk,st} = \sigma_{js}\sigma_{kt} + \sigma_{jt}\sigma_{ks}$ for $\{jk, st\} \in \{1, \ldots, p(p-1)/2\}$ by Isserlis's theorem (Isserlis, 1918).

In our numerical studies, we consider both non-sparse and sparse covariance matrices, which translates to non-sparse and sparse ceFC matrices, respectively. To generate $\boldsymbol{\Sigma}$ that is realistic in the context of neuroimaging, we calculate the sample covariance matrix for each scan/subject of the Midnight Scan Club resting-state

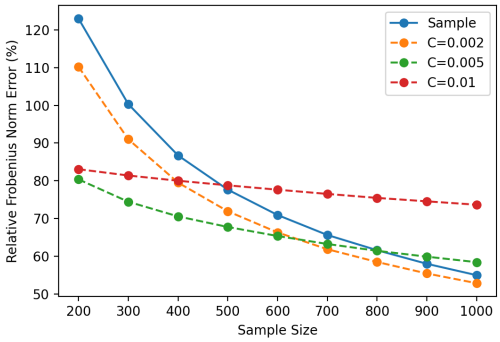
functional MRI data, and set $\mathbf{\Sigma}$ to be the average of the computed sample covariance matrices (Gordon et al., 2017). This yields a $\mathbf{\Sigma} \in \mathbb{R}^{200\times 200}$ matrix with 200 brain regions that are further partitioned into 16 brain systems (Yeo et al., 2011). A more detailed description of the MSC data set can be found in Section 2.4. To illustrate a scenario with a smaller number of variables, a straightforward situation with sufficient samples, we randomly select 25% of the brain parcels from each system, resulting in a subset of $p = 50$ parcels. The sample size n depends on the number of time frames in the MSC scans, in which the volumes are acquired approximately every 2.2 seconds. For the sparse scenario, we truncate the absolute values of $\mathbf{\Sigma}$ below 0.3 to 0, resulting in 97% sparsity for the 50×50 covariance matrix and 92% sparsity for the 200×200 covariance matrix. Finally, to ensure that the sparse covariance matrix $\mathbf{\Sigma}$ is positive definite, we set $\Sigma_{jj} = |\Lambda_{\min}(\mathbf{\Sigma})| + 0.01$ for $j = 1, \ldots, p$, where $\Lambda_{\min}(\mathbf{\Sigma})$ is the minimum eigenvalues of $\mathbf{\Sigma}$ (Tan et al., 2021). We then rescale $\mathbf{\Sigma}$ to obtain a correlation matrix such that all variables are on the same scale with variance one.

In Figure 2.1, we summarize the relative average estimation error under the Frobenius norm across 100 independent replications for the non-sparse and sparse scenarios, respectively. From Figure 2.1, we observe that estimation errors for both non-thresholded and thresholded estimators decrease as we increase the number of time frames n. Our results also suggest that the thresholded estimator can outperform the sample moment estimator when the signal-to-noise ratio is small (i.e., when n is small), even when the true ceFC matrix is dense. Under the scenario in which the ceFC matrix is sparse, we see that the proposed thresholded estimator consistently outperforms the sample moment estimator for all truncation parameter choices across varying C.

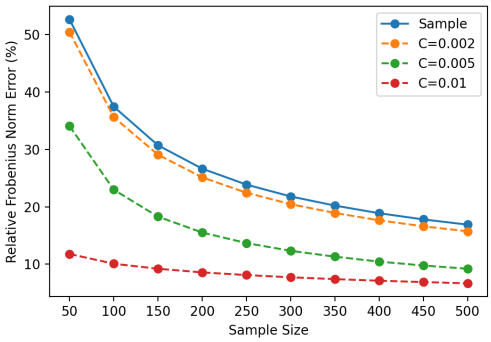
We further evaluate the thresholded estimator's performance in discerning zero and non-zero entries in a sparse ceFC matrix in terms of the true positive rate (TPR) (2.3.1) and the false positive rate (FPR) (2.3.2), calculated as the proportion of correctly estimated non-zero entries and incorrectly estimated non-zero entries, respectively. We obtain the receiver operating characteristic (ROC) curves in Figure 2.2 by varying the thresholding tuning parameter from $C = 0.0001$ to $C = 1$. The true and false positive rates, averaged over 100 scans, for $n \in \{50, 100, 200, 500, 1000\}$ and $p \in \{50, 200\}$ are summarized in Figure 2.2. We see from Figure 2.2 that the area under the curve increases as the sample size n increases, suggesting that the thresholded estimator is able to estimate the non-zero and zero entries of the true ceFC matrix more accurately as the signal-to-noise ratio increases.



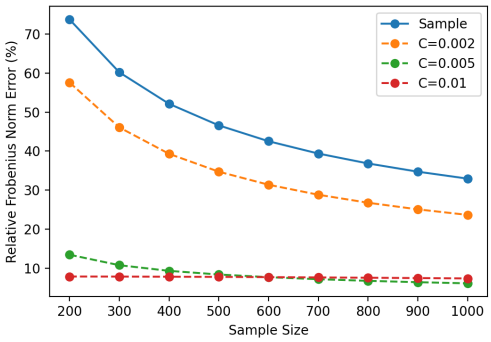
(a) Non-sparse ceFC with $p = 50$



(b) Non-sparse ceFC with $p = 200$

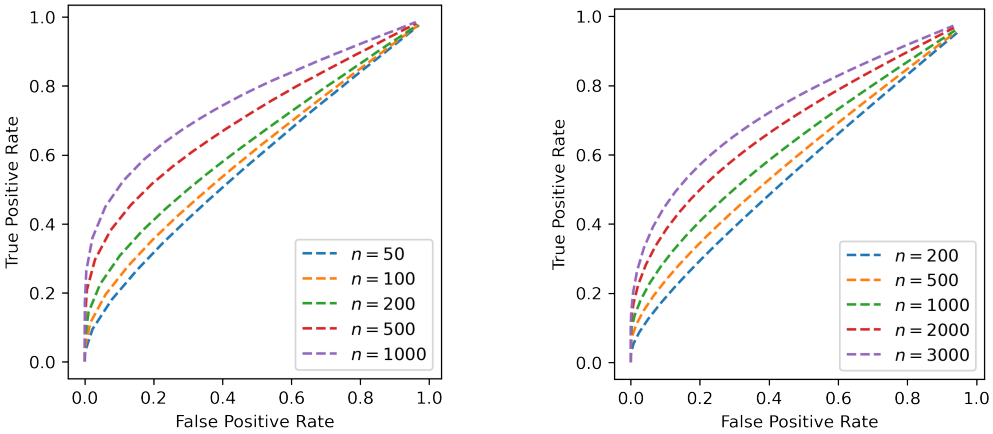


(c) Sparse ceFC with $p = 50$



(d) Sparse ceFC with $p = 200$

Figure 2.1: Relative Frobenius norm estimation error for the non-sparse (top) and sparse (bottom) ceFC matrix with the number of brain parcels set to $p = 50$ (left) or $p = 200$ (right). The colored lines correspond to the proposed sample moment estimator (—), the thresholded estimators using the tuning parameter $C = 0.002$ (– –), $C = 0.005$ (– –), and $C = 0.01$ (– –). The sample size n (number of time frames), number of brain parcels p, and tuning parameter C determine the thresholding τ through $\tau = C\sqrt{\log^4(TN^4)/T}$ for the thresholded estimator.



(a) $p = 50$ (b) $p = 200$

Figure 2.2: The true and false positive rates for sparse ceFC matrix with different choices of tuning parameter $C = 0.0001$ to $C = 1$. The colored curves correspond to the proposed thresholded estimator under different sample sizes n.

2.3.2 Comparison with eFC

As discussed in Section 2.2.5, one major distinction between eFC and ceFC is the centering step in ceFC. In this section, we empirically illustrate the impact of this centering step through directly comparing the value distributions of the eFC and ceFC matrices using the sample moment estimator from above simulation with $p = 200$ brain parcels. We visualize entry values in eFC, ceFC, and their difference in Figure A.4. In general, eFC and ceFC exhibit similar patterns in that both their mean values (eFC: 0.0019, ceFC: 0.0016) and standard deviations (eFC: 0.073, ceFC: 0.065) are close, and their difference ranges from -0.24 to 0.33. Due to the large sample size and small standard deviations, a standard t-test is overly sensitive to negligible differences and is not informative. Consequently, we instead compare the element-wise estimation bias from the upper triangular entries of eFC and ceFC, both of which are symmetric matrices with unit diagonals. To quantify the extent of overestimation relative to the ground truth Θ, we compute the overestimation ratio for each off-diagonal entry in the estimators where the estimator's absolute value exceeds the corresponding true absolute value. Specifically, for such entries, we define the overestimation ratio as $(|\widehat{\Theta}_{jk,st}| - |\Theta_{jk,st}|)/|\Theta_{jk,st}| \times 100\%$. This ratio captures the relative magnitude of overestimation with respect to the true value. As illustrated in Figure A.4(d), the distribution of these ratios, plotted on a logarithmic scale, reveals that the eFC estimator exhibits systematically higher overestimation compared to ceFC, underscoring the biased estimation in eFC. This analysis provides empirical evidence that the centering step, i.e., $\sigma_{jk}\sigma_{st}$ term, is non-trivial and essential for

accuracy. The consistently lower overestimation ratios of ceFC validate our theoretical concern that omitting this term introduces bias and leads to a less reliable estimation.

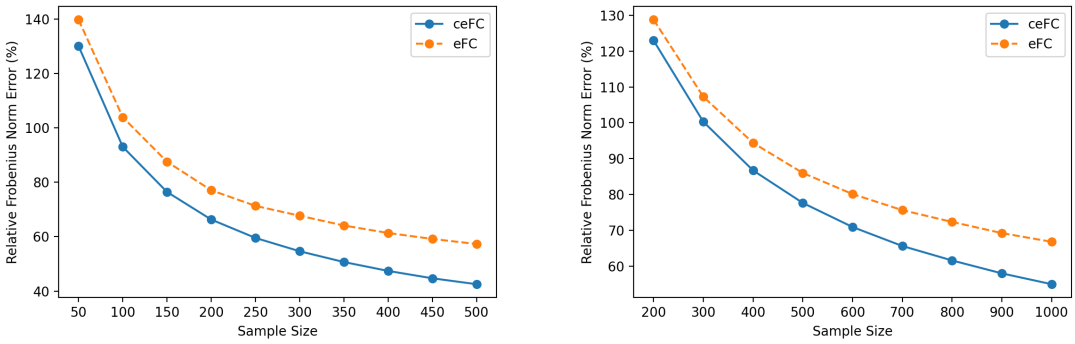
When the parameter of interest is the unbiased covariation between edge pairs, ceFC provides superior estimation accuracy, as demonstrated by the estimation error under the Frobenius norm. Under the same simulation setting, both eFC and ceFC based on sample moment estimators show decreasing estimation errors with increasing sample size, but ceFC consistently achieves smaller errors that converge to zero at a faster rate as illustrated in Figure 2.3.

Beyond estimation accuracy, it is crucial to assess a method's empirical robustness. To formally compare the robustness of eFC and ceFC, we design a split-half reliability analysis across 100 simulation trials. The primary goal of this analysis is to quantify how reliably each method could estimate higher-order connectivity. Based on a similar setting mentioned above, we split the time series into first and second halves, with $n = 500$ for each half. The eFC and ceFC matrices are computed separately for each half of the data. To determine the reliability for each edge pair, we aggregate the results across 100 trials to form two vectors, where the first vector contains the estimated eFC or ceFC values from the first half, and the second vector contains the corresponding values from the second half. The reliability score for that specific edge pair is then defined as the correlation between these two vectors. This process yields a full distribution of reliability scores for all edge pairs, allowing for a direct comparison of estimation robustness between eFC and ceFC. The results are summarized in Figure A.5, which displays the overall distributions of split-half reliability scores across all edge pairs for $p = 50$ and $p = 200$ parcels. Overall, ceFC and eFC demonstrate comparable reliability distributions. However, closer inspection reveals that ceFC shows a marginally higher count of edge pairs with positive split-half correlations, whereas eFC yields relatively more edge pairs with negative correlations. Although these distributional differences are visually indistinguishable by the large count scale in the plot, given the total number of edges considered, even minor proportional shifts can correspond to a substantial number of connections, suggesting potentially methodological distinctions in reliability behavior.

2.3.3 Hypotheses Testing on ceFC

In this section, we perform numerical studies to assess the finite-sample performance of the asymptotic normality results in Section 2.2.1. Specifically, we evaluate the proposed inference framework by testing the off-diagonal entries of a sparse ceFC matrix. Our inference procedure tests the hypothesis

$$H_{0l} : \Theta^*_{jk,st} = 0 \ v.s. \ H_{1l} : \Theta^*_{jk,st} \neq 0$$



(a) $p = 50$ (b) $p = 200$

Figure 2.3: Relative Frobenius norm estimation error for a non-sparse ceFC matrix. The colored lines correspond to the proposed sample moment estimation for ceFC (—) and eFC (– –).

for each l^{th} off-diagonal entry $\Theta^*_{jk,st}$ where Θ^* is the true ceFC matrix. That is, we examine all possible combinations of edges to see whether each pair, denoted as jk, st , exhibits a strong association simultaneously. The test statistic $\gamma_{jk,st}$ is constructed as defined in (2.2.4) for each hypothesis, which is asymptotically normal as shown in (2.2.3).

The simulation setting for the inference framework is similar to the process in 2.3.1, except that we generate a sparse ceFC matrix with moderate signals. To achieve this, we adopt the approach in Tan et al. (2021) by generating a sparse covariance matrix with some non-zero values. Specifically, we randomly assign 0.9 to 5% and 2% of off-diagonal entries in an identity matrix for the $p = 50$ and $p = 200$ cases, respectively. We then employ the same procedure as mentioned earlier to ensure the resulting sparse covariance matrix remains positive definite.

To evaluate the numerical performance of the inferential framework in the finite sample setting, we calculate type I error and power as the proportion of falsely rejected and correctly rejected null hypotheses, respectively. Based on the data generation scheme, the centered edge connectivity matrices include both weak and moderate signals. To this end, we consider the inference for non-zero entries in two scenarios, one for moderate signals and one for weak signals that may require more samples for detection. Based on the distribution of non-zero values in the constructed ceFC matrix, we consider values within $(0, 0.2]$ as weak signals and values within $(0.2, 0.5]$ as moderate signals. For computational purposes, we randomly select 1000 zero entries from the upper-triangular of the ceFC matrix for type I error, and 1000 weak signals and 1000 moderate signals for power study, respectively. This process of random selection is repeated 5000 times. Specifically, for each of these 5000 iterations, we

calculate the type I error based on the 1000 zero entries and the statistical power based on the 1000 non-zero entries (weak and moderate signals), respectively. The type I error and power are then averaged over these 5000 iterations and reported in Table 2.1. From Table 2.1, we can see that the type I error is well controlled at the 0.05 significance level, and the power increases to one as the sample size increases, while we need a larger number of samples for the power of weak signals to be expected.

Table 2.1: The type I error and power for testing the 1000 random off-diagonal entries in ceFC at the 0.05 significance level are averaged over 5000 data sets. The power study is separated into 1000 random weak signals $\in (0, 0.2]$ and 1000 random medium signals $\in (0.2, 0.5]$, respectively. Simulation results for $p = 50$ and $p = 200$ variables over a range of sample size n are reported.

	$p = 50$			$p = 200$		
	Type I Error	Weak Signal	Medium Signal	Type I Error	Weak Signal	Medium Signal
$T = 300$	0.048	0.237	0.858	0.048	0.163	0.755
$T = 500$	0.049	0.377	0.971	0.049	0.250	0.922
$T = 800$	0.049	0.554	0.998	0.049	0.377	0.988
$T = 1000$	0.049	0.650	0.999	0.049	0.455	0.997
$T = 1500$	0.050	0.813	1	0.049	0.621	1
$T = 2000$	0.050	0.904	1	0.050	0.746	1
$T = 3000$	0.050	0.972	1	0.050	0.894	1
$T = 4000$	0.050	0.989	1	0.050	0.958	1
$T = 5000$	0.050	0.994	1	0.050	0.983	1
$T = 6000$	0.050	0.997	1	0.050	0.993	1
$T = 7000$	0.050	0.998	1	0.050	0.997	1
$T = 8000$	0.050	0.999	1	0.050	0.998	1
$T = 10000$	0.050	1	1	0.050	0.999	1
$T = 15000$	0.050	1	1	0.050	1	1

We further investigate the multiple hypothesis testing framework with FDR, which is computed as the proportion of falsely rejected null hypotheses among all rejections. Under the same data generation scheme described above, we choose a random subset of upper-triangular entries from the ceFC matrix. Specifically, for $p = 50$ (i.e., 50 brain parcels), we select 3000 entries, and for $p = 200$, we select 5000 entries. These subsets of entries consist of both zero and non-zero values reflecting the sparsity in the true ceFC matrix. We summarize FDR with significance level $\alpha = \{0.05, 0.1, 0.2\}$ over 5000 data sets in Table 2.2. As our multiple hypothesis testing procedure is based on the procedure of Benjamini and Yekutieli (2001), our FDR results can be conservative, as illustrated in Table 2.2. The results in Table 2.2 show that the FDR is well controlled under each significance level, which is consistent with the theoretical

result.

Table 2.2: The false discovery rate for testing 3000 random entries for $p = 50$ and 5000 random entries for $p = 200$ are averaged over 5000 data sets. These random subsets retain the sparsity in the corresponding ceFC matrix. The significance level α is set to be $\{0.05, 0.1, 0.2\}$ and the sample size n ranges from 400 to 700.

p	α	$n = 400$	$n = 450$	$n = 500$	$n = 550$	$n = 600$	$n = 650$	$n = 700$
$p = 50$	$\alpha = 0.05$	0.017	0.012	0.007	0.005	0.004	0.004	0.005
	$\alpha = 0.1$	0.020	0.015	0.011	0.009	0.009	0.009	0.011
	$\alpha = 0.2$	0.027	0.023	0.019	0.017	0.018	0.019	0.021
$p = 200$	$\alpha = 0.05$	0.057	0.035	0.019	0.010	0.008	0.006	0.005
	$\alpha = 0.1$	0.060	0.035	0.022	0.015	0.012	0.009	0.009
	$\alpha = 0.2$	0.066	0.039	0.028	0.021	0.019	0.016	0.017

2.4 An Application to the Midnight Scan Club Data

A key feature of the edge-centric framework is that it allows direct and natural identification of overlapping communities. Substantial evidence supports the existence of overlapping structures in brain networks (Palla et al., 2005; Yeo et al., 2014; Beckmann et al., 2005; van den Heuvel and Sporns, 2013), partly motivated by responses to the extensive body of literature studying non-overlapping communities (Power et al., 2011a; Yeo et al., 2011). In this section, we demonstrate how the proposed estimators for ceFC can be used to identify overlapping communities and provide deeper insights into brain network organization (Faskowitz et al., 2020; Chumin et al., 2022), which can be potentially developed as biomarkers of various disorders (Dai et al., 2023; Yang et al., 2023; Zamani Esfahlani et al., 2022; Sun et al., 2023).

We apply our proposed methods to the MSC data (Gordon et al., 2017) with $p = 200$ brain regions of interest. The MSC data (Gordon et al., 2017) consists of fMRI images of 10 healthy adults (5 females, age: $24-34$ years old). Informed consent was obtained from all participants, and the study received approval from the Human Studies Committee and Institutional Review Board at Washington University School of Medicine. Each participant underwent 10 30-minute scan sessions with eyes open on separate days, with each scan starting from midnight. All images, including four T1-weighted and four T2-weighted images, were collected by a Siemens TRIO 3T MRI scanner. For a detailed description of data acquisition, please see Gordon et al. (2017).

These fMRI measurements were pre-processed with slice-time correction, motion

correction, intensity normalization, atlas transformation with registration to T2 image, and distortion correction (Gordon et al., 2017). Subsequently, the pre-processed fMRI data underwent demeaning and detrending processes, nuisance regression, high-motion frame removal with framewise displacement (FD) over 0.2 mm, and bandpass filtering between 0.009 Hz and 0.08 Hz (Power et al., 2014). These functional measurements were further transposed to the cortical surface with a two-dimensional geodesic smoothing.

The pre-processed resting-state fMRI measurements were aggregated within each subject, resulting in 4000 to 8000 time frames for 10 subjects, i.e. T ranges from 4000 to 8000. These BOLD signals were mapped onto the Schaefer2018 parcellation (Schaefer et al., 2018), defining 200 brain regions of interest, or equivalently $p = 200$. These brain regions can be further grouped into canonical functional networks based on the system-level structure proposed by Yeo et al. (2011).

In Section 2.4.1, we estimate the ceFC matrix for the $p = 200$ brain regions of interest using the proposed methods. We then perform clustering on the resulting ceFC matrix to identify overlapping communities of the brain regions. We also perform a sensitivity analysis to assess the stability of the proposed methods in Section 2.4.2.

2.4.1 Identifying Overlapping Communities

We perform spectral clustering on the estimated ceFC matrix to identify overlapping communities of the brain regions. Specifically, we estimate the ceFC for each subject using (i) the sample moment estimator in (3.2.3) (ii) the FDR-controlled estimator, obtained by testing whether each entry of the computed sample moment estimator is equal to zero with the FDR controlled at a level of 0.1, and zeroing insignificant entries; and (iii) the thresholded estimator in (2.2.5).

The resulting FDR-controlled estimators exhibit slightly varying degrees of sparsity, with levels ranging from 0.6 to 0.8 across all subjects. The thresholded estimator (2.2.5) involves selecting a tuning parameter τ–we select τ such that the sparsity level of the thresholded estimator aligns with that of the FDR-controlled estimator for each subject. Finally, we compute the average for each of the three estimators across 10 subjects.

After obtaining edge communities by spectral clustering, we project them back to the brain region level following the procedure in Faskowitz et al. (2020) to obtain overlapping communities for the brain regions. Motivated by Faskowitz et al. (2020), we perform spectral clustering on each of the aforementioned three ceFC estimators to identify overlapping communities of the brain regions. That is, we conduct eigendecomposition on each of the ceFC estimators, extract 50 eigenvectors with the

highest eigenvalues, and subsequently perform $k-$means clustering with $k = 5$ on the resulting eigenvectors with 5 communities for visualization purposes.

To obtain overlapping communities for the brain regions, we project the edge communities back to the brain region level following the procedure in Faskowitz et al. (2020). Specifically, for each brain region, we calculate the proportion of edges in each of the estimated 5 communities, allowing each brain region to belong to multiple communities with different weights. We note that cluster labels from the three different estimators can be different. To address this issue, we relabel the cluster labels by maximizing the pairwise Matthews correlation coefficients between a pair of estimators (Matthews, 1975). That is, we relabel the community labels for each estimator using the above method by matching them to the average thresholded estimator.

The resultant communities are presented in Figures 2.4a,2.4b, and 2.4c, representing results based on the sample moment estimator, the FDR-controlled estimator, and the thresholded estimator, respectively.

These results in general show that the three different estimators reveal similar community structures (ANOVA: $F(2, 2997) = 0.1$, $p = 0.9$. Each observation represents the proportion of edges assigned to a specific community for a given brain region as shown in Figures 2.4a,2.4b, and 2.4c, with 200 brain regions $\times$ 5 communities per estimator). Specifically, the sample moment estimator, as shown in the Figure 2.4a, predominantly identifies communities 1, 2, and 4, with the average proportions across all brain parcels as $25.7\%, 31.5\%$, and 29.3%, respectively. In contrast, both FDR-controlled (Figure 2.4b) and thresholded estimators ((Figure 2.4c) additionally group the brain regions into communities 1, 2, 4, and 5 with similar proportions. Specifically, the FDR-controlled estimator shows average proportions for these four communities as $29.8\%, 26.9\%, 23\%$, and 11.5%, respectively, and the thresholded estimator returns similar outcomes as $31.8\%, 28.3\%, 20.8\%$, and 10.7%. The similarity across the three methods can be further quantified through the correlation of proportions of edges within each community between each pair of methods. As shown in Figure 2.4d, all three methods present highly correlated community structures, with more similar results between the FDR-controlled and thresholded estimators, demonstrating that the identified community structures are robust to the specific statistical pipeline applied to the ceFC results. We further examine the significance of these correlations against a null model using a spin permutation test. Specifically, we randomly shuffle the 200 brain parcels 1000 times while preserving their spatial autocorrelation structure. The corresponding correlation is then computed for each shuffle to generate a null distribution. The results are summarized in Figure 2.4e, where the observed correlations for all comparisons were significantly higher than the values generated from the permuted data ($p < 0.001$ from 1000 data points or shuffles), suggesting these similarities

are significantly higher than what is expected given the spatial autocorrelation of the fMRI data when the indices are shuffled.

The projection patterns indicate varying involvement of brain systems across different communities. Specifically, community 1 is dominated by default and limbic networks; community 2 is primarily influenced by default, control, and temporoparietal networks; community 3 is mostly composed of somatomotor and attention networks; community 4 mainly comprises control and attention networks; and community 5 is primarily associated with attention, somatomotor, and visual networks. Note, however, that nodal memberships in each community are graded. For instance, community 2 from the sample moment estimator (Figure 2.4a) is, indeed, dominated by default mode regions – i.e., those regions are most strongly affiliated with this community – but even somatomotor regions exhibit weak, but non-negligible, affiliation. To systematically assess the significant proportion differences among these brain systems, we first identified the presence of any disparities among brain systems within each community and then applied Tukey's HSD test for multiple comparisons between each pair of brain systems, utilizing adjusted p-values at a significance level of 0.05. For each system, we compiled the number of outcomes where a system showed significantly greater proportions than others and summarized these findings in Figure A.1, reporting results for each community and method. The top three systems with the highest counts of significant outcomes are highlighted in Figure 2.5. Here, $***$ indicates the system with the highest number of significant outcomes, $**$ denotes the second highest, and $*$ signifies the third highest, reflecting the levels of significance of the proportions within each system.

To evaluate the overlapping structures across communities, we compute the normalized entropy using the proportions across the 5 communities within each brain system. This metric measures the uncertainty of community delineations: values closer to 1 indicate greater overlap across communities, while values closer to 0 suggest a more concentrated distribution within a single community. The results, displayed in Figure A.2, show higher entropy values in the attention and somatomotor networks. On the surface, these findings of substantial functional overlap in these specific networks are consistent with findings from a previous study (Faskowitz et al., 2020). However, our ceFC framework provides a more robust interpretation of this result by addressing a critical source of bias in the eFC definition that could impact reproducibility. Mathematically, eFC is sensitive to the mean of edge time series (i.e., static functional connectivity), introducing a sample-specific bias. In contrast, our proposed ceFC is, by definition, a centered covariation that explicitly removes such bias.

The similarity of our results to those from eFC, despite the removal of the con-

founding bias, suggests that while the original eFC findings were directionally correct, they possess an inherent vulnerability that their reproducibility in future studies would depend on new datasets sharing the same static connectivity biases. Our ceFC results, therefore, are not a simple replication. Instead, they serve as a crucial validation, isolating the true, reproducible phenomenon from a potentially biased observation.

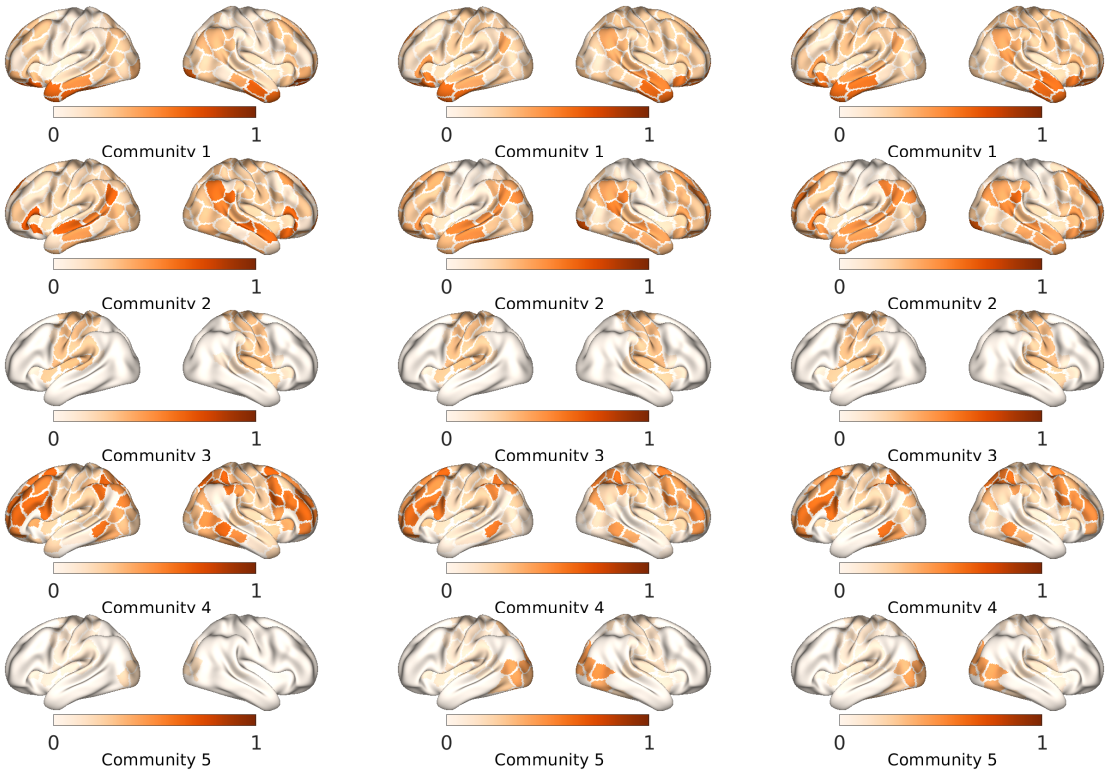
With this refined understanding, we further compare the normalized entropy between each pair of brain systems and record the number of significantly higher entropy outcomes for each system, as shown in Figure A.3. This analysis further indicates that the aforementioned brain systems, i.e., attentional and somatomotor networks, exhibit significantly greater overlapping structures.

While previous results show broadly similar patterns across the three estimators, there are certain differences among them. In contrast to the FDR-controlled and thresholded estimators, the sample moment estimator reveals subtle variances in community structures, particularly among systems associated with functional tasks. It is not surprising to observe similar results between the FDR-controlled and thresholded estimators, as the tuning parameter τ in the thresholded estimator is chosen to align its sparsity with that of the FDR-controlled estimator. Specifically, Figure 2.5 indicates that the somatomotor networks are more active in community 1 by the FDR-controlled and thresholded estimators but in community 2 by the sample moment estimator. Similarly, the visual systems are more prominent in community 5 by the FDR-controlled and thresholded estimators, but are relatively more effective in community 4 by the sample moment estimator. In contrast, the somatomotor and temporoparietal systems exhibit higher activity in community 5 by FDR-controlled and thresholded estimators, but tend to be more effective in community 4 by the sample moment estimator. Previous studies have highlighted the significant engagement of the somatomotor and visual systems in respective motor and visual functions (Biswal et al., 1995; Chenji et al., 2016; Yang et al., 2015). These distinctions underscore the value of exploring system behavior related to functional tasks within a resting-state context.

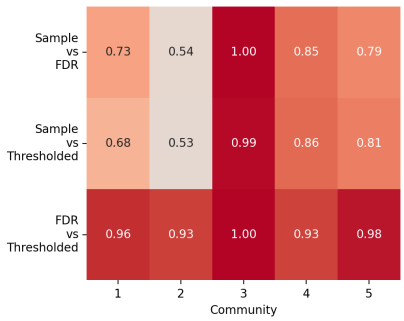
We now perform a formal statistical hypothesis test to assess whether the average proportion within each brain system significantly exceeds zero. Specifically, for each system s, we compute its average proportion μ_{cs} within each community c and test the following hypothesis

$$H_0 : \mu_{cs} = 0 \; vs. \; H_1 : \mu_{cs} > 0 \tag{2.4.1}$$

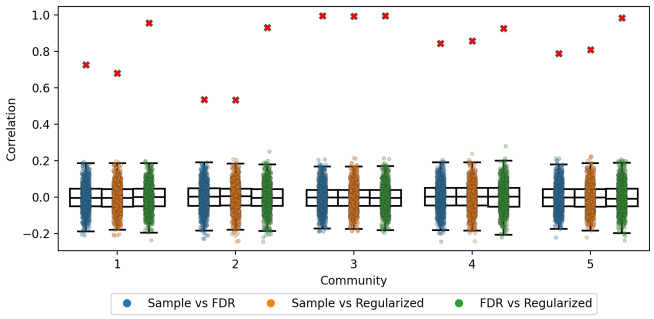
with a significance level of 0.05. While the null model may fail to preserve the



(a) Sample moment (b) FDR controlled (c) Thresholded

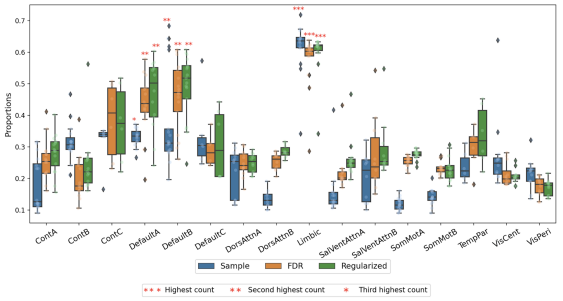


(d) Correlation of proportions of edges

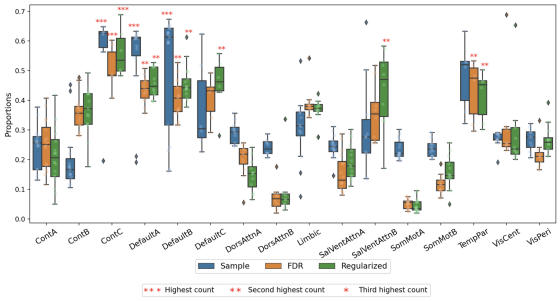


(e) Permutation

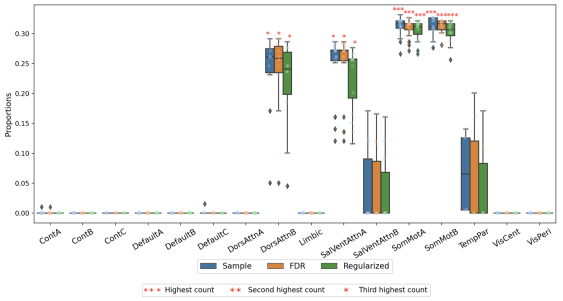
Figure 2.4: For all edges connected to each brain region, we compute the proportions of edges within each community. Panels (a)–(c) correspond to results based on the average sample moment estimator, the average FDR-controlled estimator with FDR controlled at level 0.1, and the average thresholded estimator, respectively. Panel (d) shows the correlations of the edge-proportion summaries between each pair of methods. Panel (e) shows the corresponding correlations under random permutations over 1000 iterations. The observed correlations from the original results are marked as red X's.



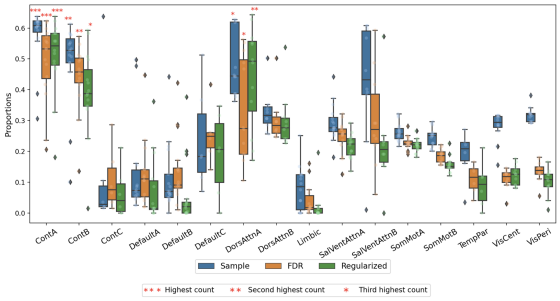
(a) Community 1



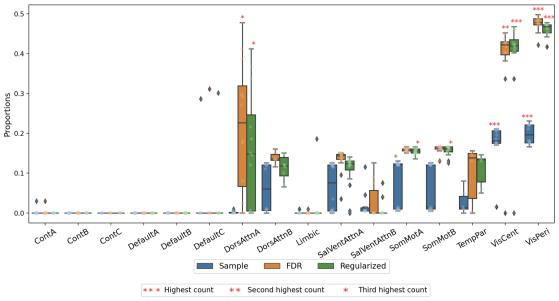
(b) Community 2



(c) Community 3

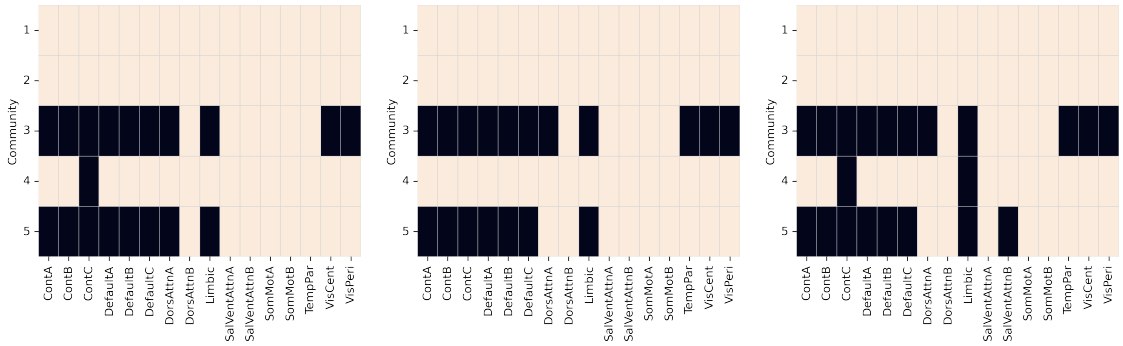


(d) Community 4



(e) Community 5

Figure 2.5: Proportions of edges within each community and associated brain systems as depicted in Figure 2.4. The plot highlights the top three brain systems that demonstrate a significantly greater proportion of edges in pairwise comparisons with other systems. $***$ denotes the system with the highest count of significant outcomes, $**$ represents the second highest, and $*$ signifies the third highest.



(a) Sample moment estimator (b) FDR-controlled estimator (c) Thresholded estimator

Figure 2.6: Brain systems with significant proportions from each average estimator under the significance level of 0.05. The black box represents the insignificant result, and the yellow box represents the significant result.

spatial structure of the original data (Váša and Mišić, 2022), our objective is to evaluate the performance of these estimators statistically. Therefore, we proceed with the same testing procedure for the proportions from each subject-specific estimator and the average estimator, respectively. For visualization purposes, the test results from the average estimators are summarized in Figure 2.6. The results in Figure 2.6 indicate general consensus among the three estimators regarding the systems with significant proportions within communities, although subtle disparities are observed in communities 3, 4, and 5.

2.4.2 Stability of the Proposed Estimators

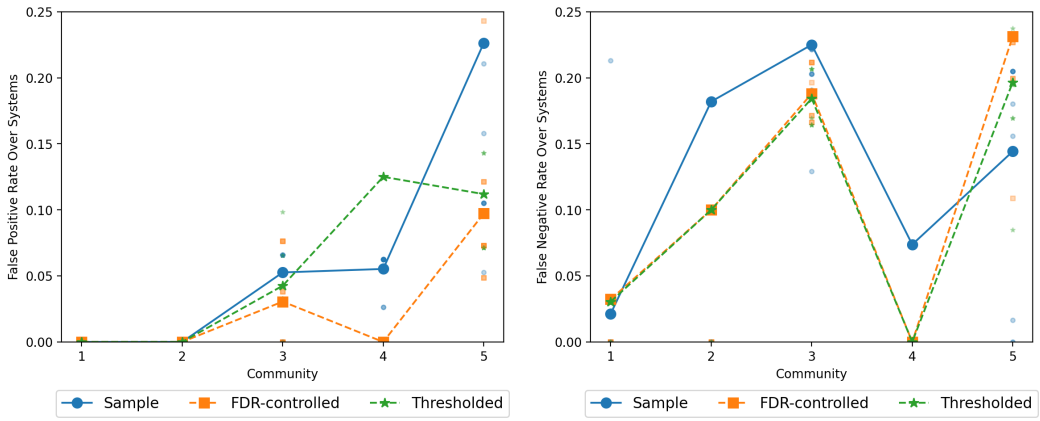
In this section, we conduct an extensive sensitivity analysis to evaluate the stability of the three estimators across different individuals, based on the hypothesis testing (2.4.1) on the community proportions obtained from each type of estimator as described in Section 2.4.1. We consider the test results from each average estimator, denoted as μ^{avg}, as ground truth. Across the subject-specific test outcomes, denoted as μ^{sub}, we define the false positive rate (FPR) and the false negative rate (FNR) as the proportion of communities c or systems s incorrectly identified as significant and the proportion of communities c or systems s incorrectly identified as insignificant across all communities, respectively:

$$\begin{aligned} \text{FPR}(\mu^{\text{sub}}, \mu^{\text{avg}}) &= \frac{\#\{(c,s) : \mu_{cs}^{\text{sub}} > 0 \text{ and } \mu_{cs}^{\text{avg}} = 0\}}{\#\{(c,s) : \mu_{cs}^{\text{avg}} = 0\}}; \\ \text{FNR}(\mu^{\text{sub}}, \mu^{\text{avg}}) &= \frac{\#\{(c,s) : \mu_{cs}^{\text{sub}} = 0 \text{ and } \mu_{cs}^{\text{avg}} > 0\}}{\#\{(c,s) : \mu_{cs}^{\text{avg}} > 0\}}. \end{aligned} \tag{2.4.2}$$

The results averaged over 10 subjects are reported in Table 2.3. It can be seen from Table 2.3 that both FDR-controlled and regularized estimators achieve lower false positive rates and false negative rates compared to the results from the sample moment estimator. Specifically, the false positive and false negative rates are reduced by 61% and 15% in the FDR-controlled estimator and reduced by 16% and 21% in the thresholded estimator, respectively. Additionally, we compute the average false positive and average false negative rates for all the three methods at the community and brain system levels, respectively. The results are displayed in Figure 2.7. We see from Figure 2.7 that the sample moment estimator generally demonstrates comparable or higher false positive and false negative rates across the majority of communities and brain systems. The FDR-controlled and regularized estimators still present similar false negative rates, while the regularized estimator returns slightly higher false positive rates in certain communities and brain systems. Our results suggest that the proposed FDR-controlled and regularized estimators, which disregard the noise and weak signals in edge connectivity, are likely to yield more stable outcomes than the sample moment estimator.

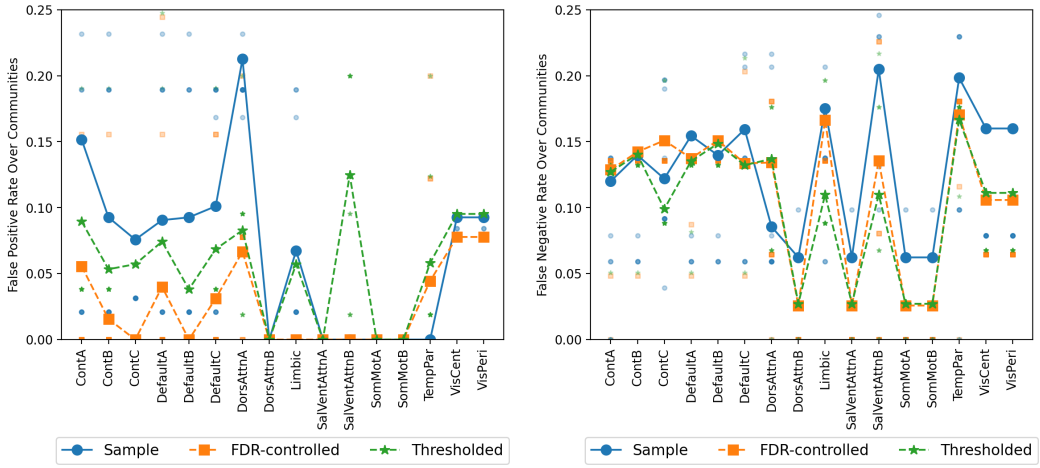
Table 2.3: False positive rate (FPR) and false negative rate (FNR) are averaged across 16 brain systems and 5 communities over 10 subject data for each estimator.

Method	FPR	FNR
Sample Moment	0.067	0.129
FDR-controlled	0.026	0.110
Regularized	0.056	0.102



(a) False Positive Rate Averaged Over Systems

(b) False Negative Rate Averaged Over Systems



(c) False Positive Rate Averaged Over Communities

(d) False Negative Rate Averaged Over Communities

Figure 2.7: The false positive and false negative rates are averaged across 16 systems (first row) or 5 communities (second row). The solid points on line graphs represent the average results over subjects, and each small transparent point corresponds to each subject's result.

CHAPTER 3

Assessing the Predictive Value of Centered Edge Functional Connectivity for Phenotypic Traits

3.1 Introduction

Chapter 2 has introduced centered edge functional connectivity (ceFC) as an unbiased higher-order measure of covariation among brain edges and established estimation and inference procedures for this representation. The results in Chapter 2 show that ceFC can recover meaningful higher-order edge relationships and reveal overlapping community structure beyond conventional node-centric summaries. Building on this foundation, it is important to examine whether such higher-order connectivity also contains useful predictive information of human, such as phenotypic traits.

This question is closely related to brain-wide association studies (BWAS), where large-scale connectivity features are used to explain or predict subject-level outcomes. A central objective in this literature is to understand how variation in brain connectivity relates to variation in behavior and phenotype across individuals (Zatorre et al., 2012; Rosenberg et al., 2018; Cheng et al., 2015; Krakauer et al., 2017; Ooi et al., 2022b). Existing studies have used connectivity features to investigate a broad range of phenotypic traits, including age (Geerligs et al., 2015; Varangis et al., 2019), motion (Fox et al., 2007), cognition (Ooi et al., 2022a; Chen et al., 2022), intelligence (Van Den Heuvel et al., 2009), memory (Wang et al., 2010; Hampson et al., 2006), personality (Fulwiler et al., 2012; LaBar et al., 2001), and clinical outcomes (Marek et al., 2022; Rosenberg and Finn, 2022; Gratton et al., 2022; Pan et al., 2024; Yahata et al., 2016).

To examine these brain-phenotype associations, researchers have increasingly turned to multivariate predictive models. Representative approaches include support vector regression (Marek et al., 2022), canonical correlation analysis (Smith et al., 2015; Marek et al., 2022), and partial least squares (Helmer et al., 2024; Krishnan et al., 2011). The performance of these models is typically evaluated by comparing predicted and observed phenotypic values using metrics such as correlation and mean

squared error, which provide measures of predictive accuracy (Smith et al., 2015; Marek et al., 2022; Kang et al., 2024).

Most predictive analyses in this setting are based on conventional node-centric functional connectivity (nFC). By contrast, ceFC provides a higher-order edge-centric representation of brain connectivity. Such higher-order structure may capture interaction patterns that are not directly available from conventional nFC and may therefore improve brain-phenotype prediction. On the other hand, ceFC is mainly built on the basis of nFC but with a much higher dimensionality and may still reflect variation that is already largely contained in lower-order nFC structure. Therefore, it is important to analyze whether the increased representational complexity leads to improved predictive performance or mainly re-expresses information already captured by conventional nFC.

To address this issue, this chapter adopts a standard BWAS framework and constructs subject-level representations based on both nFC and ceFC in order to compare their predictive utility for a range of phenotypic traits. Specifically, the nFC measurements are organized into a subject-by-node-pair matrix $\boldsymbol{R}_{\text{full}}$ by vectorizing the upper triangular entries of each subject's nFC matrix and stacking the resulting vectors across subjects. Because the full ceFC matrix is high-dimensional, we compute regularized ceFC estimates introduced in Chapter 2 for each subject and reduce the dimension by retaining only the non-weak edge pairs across all subjects, yielding a reduced subject-by-edge-pair matrix $\boldsymbol{\Theta}_{\tau}$. We then extract principal components (PCs) from both $\boldsymbol{R}_{\text{full}}$ and $\boldsymbol{\Theta}_{\tau}$, referred to as nFC PCs and ceFC PCs, respectively. These PCs capture the dominant modes of variation in nFC and ceFC across subjects and serve as candidate predictors in subsequent analyses.

For a range of phenotypic traits as response variables, we construct baseline linear regression models and more advanced partial least squares (PLS) models using various combinations of the nFC and ceFC PCs, including (i) nFC PCs alone, and (ii) nFC PCs together with ceFC PCs. Model performance is evaluated based on the correlation and mean squared error between predicted and observed phenotypic values. We find that, when a sufficiently rich set of nFC PCs is included, integrating ceFC PCs as additional predictors does not meaningfully improve predictive performance. This result suggests that ceFC PCs contain limited additional information beyond what is already captured by nFC for the traits considered here. To better understand this finding, we regress each leading ceFC PC on the full set of nFC PCs and obtain high R^2 values, indicating that much of the variation in ceFC PCs can already be explained by existing nFC components. In other words, the dominant structures represented by ceFC appear to be largely redundant with those already captured in conventional nFC, thus offering little additional predictive information

for the phenotypic traits examined in this study.

Taken together, this chapter evaluates whether the proposed ceFC in Chapter 2 improves the prediction of phenotypic traits beyond conventional node-centric FC within a BWAS framework. In addition to comparing predictive performance across modeling strategies, our analysis also examines the relationship between ceFC-derived components and conventional nFC-derived components in order to clarify the source of the observed redundancy. These results help refine our understanding of the role of higher-order connectivity in brain-phenotype prediction. More broadly, they suggest that increased representational complexity does not necessarily translate into improved predictive power, especially when much of the variation is already captured by lower-order functional connectivity features.

3.2 Methods

3.2.1 AOMIC ID1000 Data

In this study, we use ID1000, a multimodal 3T MRI dataset from the Amsterdam Open MRI Collection (AOMIC), to construct the nFC and ceFC patterns described in the following section. The ID1000 dataset contains fMRI scans obtained from 928 healthy individuals aged between 19 and 26, with an educational profile comparable to that of the general Dutch population (Snoek et al., 2020). Limiting the sample to this age range helps reduce the potential influence of aging on brain-related measures. The study protocol was approved by the faculty's ethics committee, and all participants provided informed consent before scanning.

During data acquisition, the ID1000 set collected three T1-weighted structural scans, three diffusion-weighted scans, and one functional BOLD MRI scan through a Philips 3T scanner (Snoek et al., 2020). In this chapter, we use the fMRI scan from each subject, which recorded the time series data for node signals, to construct nFC and ceFC. Each fMRI scan lasted approximately 11 minutes, during which participants viewed a movie clip composed of 22 natural scenes accompanied by music (Reggio and Coppola, 1982). These scenes cover a range of visual conditions involving different textures and objects with different sizes and movements. The duration of individual scenes varied from approximately 5 to 40 seconds. These fMRI scans were preprocessed with motion correction, field distortion correction, and BOLD-to-T1w transformation (Snoek et al., 2020). We then applied the Schaefer2018 parcellation (Schaefer et al., 2018) to the preprocessed fMRI data and obtained 200 brain regions for subsequent analysis.

In addition to the fMRI scans, the ID1000 dataset contains detailed demographic

and psychometric information. In particular, participants were asked to fill out a set of questionnaires covering general demographic characteristics and personality-related measures (Snoek et al., 2020). These include age, body mass index (BMI), the Intelligence Structure Test (IST) (Liepmann et al., 2010; Amthauer et al., 2001), which evaluates fluid intelligence, crystallized intelligence, and memory through verbal, numerical, and figural tasks; the State-Trait Anxiety Inventory (STAI) (Spielberger et al., 1971; Van der Ploeg, 1980), and the Behavioral Inhibition System/Behavioral Activation System (BIS/BAS) questionnaire (Carver and White, 1994; Franken et al., 2005), which measures sensitivity to punishment and reward. In this study, we use 20 numerical outcomes as the set of response variables in the prediction task.

3.2.2 Node-centric Functional Connectivity (nFC)

Functional connectivity has been widely employed in neuroimaging studies to quantify the dependence between pairs of brain parcels or nodes based on their BOLD signal time series from fMRI scans (Friston et al., 1993; Aertsen et al., 1989; Friston, 1994), providing insights into the brain's functional organization and serving as biomarkers for a range of diseases (Greicius et al., 2003; Wen et al., 2018; Yeo et al., 2011). As discussed in Chapter 2, one of the most common measures of node-centric functional connectivity (nFC) is the Pearson correlation between the time series of two brain nodes (Friston et al., 1993).

In this chapter, after applying the Schaefer2018 parcellation, each subject is represented by $p = 200$ brain regions, and each region is associated with a standardized BOLD time series of length n. Let $\boldsymbol{X}_1^{(l)}, \ldots, \boldsymbol{X}_n^{(l)}$ be n independently and identically distributed copies of $\boldsymbol{X}^{(l)} \in \mathbb{R}^p$, which represent the regional time series matrix for subject l with n time points and p regions. For each subject, the time series for node j, $j = 1, \ldots, p$, is standardized to have mean zero and unit variance. The nFC between node j and node k for subject l can be estimated as

$$\widehat{R}_{jk}^{(l)} = \frac{1}{n} \sum_{i=1}^{n} X_{ij}^{(l)} X_{ik}^{(l)},$$

which is equivalent to the sample correlation between the two standardized regional time series. Repeating this calculation for all pairs of regions yields a subject-specific node-by-node correlation matrix $\widehat{\boldsymbol{R}}^{(l)} \in \mathbb{R}^{p \times p}$. Since $\widehat{\boldsymbol{R}}^{(l)}$ is symmetric and has ones on the diagonal, we retain only its upper triangular entries excluding the diagonal, vectorize them into a subject-level feature vector, and then stack these vectors across all subjects. This produces the full-size subject-by-node-pair matrix, denoted by $\widehat{\boldsymbol{R}}_{\text{full}}$. For the ID1000 data analyzed in this chapter, after sample filtering and cleaning,

we obtain a matrix $\widehat{\boldsymbol{R}}_{\text{full}}$ of dimension 847 × 19,900, where 19,900 = 200 × 199/2 corresponds to the total number of unique node pairs.

3.2.3 Centered Edge Functional Connectivity (ceFC)

While nFC has been widely employed to study functional organization of brain networks (Bullmore and Sporns, 2009; Power et al., 2011a; Finn et al., 2015), such low-order relationships may not be sufficient to capture more complex overlapping community structures (Palla et al., 2005; Yeo et al., 2014; Beckmann et al., 2005; van den Heuvel and Sporns, 2013), for which higher-order interactions may also play an important role (Glickfeld and Olsen, 2017; Gatica et al., 2021; Luppi et al., 2022). As discussed in Chapter 2, we extend the analysis of relationships in brain networks to centered edge-centric functional connectivity (ceFC), which quantifies the unbiased covariation between two edge time series. Our results in Chapter 2 show that this higher-order relationship helps reveal more stable overlapping brain community structures and provides additional insights into brain organization. In this chapter, we briefly recall the sample moment estimator of ceFC and describe how it is applied to the ID1000 data. Given the standardized BOLD signal $X_j^{(l)}$ for subject l and brain region $j, j = 1, \ldots, p$, the connectivity between edges $\{j, k\}$ and $\{s, t\}$ can be estimated as

$$\widehat{\Theta}_{jk,st}^{(l)} = \frac{1}{n}\sum_{i=1}^{n} X_{ij}^{(l)} X_{ik}^{(l)} X_{is}^{(l)} X_{it}^{(l)} - \frac{1}{n}\sum_{i=1}^{n} X_{ij}^{(l)} X_{ik}^{(l)} \cdot \frac{1}{n}\sum_{i=1}^{n} X_{is}^{(l)} X_{it}^{(l)}.$$

This quantity measures the covariance between the two edge time series after centering. Repeating this calculation over all pairs of edges yields a subject-specific edge-by-edge matrix $\widehat{\boldsymbol{\Theta}}^{(l)} \in \mathbb{R}^{e \times e}$, where $e = p(p-1)/2$ is the number of unique edges. Because $\widehat{\boldsymbol{\Theta}}^{(l)}$ is symmetric, we retain its upper triangular entries excluding the diagonal, vectorize them, and stack the resulting vectors across all subjects to form the full-size subject-by-edge-pair matrix $\widehat{\boldsymbol{\Theta}}_{\text{full}}$. For the ID1000 data analyzed in this chapter, $p = 200$, so $e = 19{,}900$, and the resulting matrix $\widehat{\boldsymbol{\Theta}}_{\text{full}}$ has dimension $847 \times 197{,}995{,}050$, where the number of columns represents the number of edge pairs calculated as $e(e-1)/2$.

3.2.4 Dimension Reduction

For the ID1000 data analyzed in this chapter, both the full node-centric matrix $\widehat{\boldsymbol{R}}_{\text{full}}$ and the full edge-centric matrix $\widehat{\boldsymbol{\Theta}}_{\text{full}}$ are high-dimensional, in the sense that the number of subjects is much smaller than the number of node-pair or edge-pair features. In particular, this imbalance is especially severe for $\widehat{\boldsymbol{\Theta}}_{\text{full}}$, whose dimension is far beyond

what can be directly used in downstream predictive modeling. As a result, it is not feasible to include all node or edge connectivity measures as candidate predictors for phenotypic outcomes. To address this issue, we apply a two-step dimension reduction procedure. The first step uses thresholding to regularize ceFC estimates and remove weak edge-pair interactions, and the second step further summarizes the reduced representations through principal component analysis. This procedure is designed to reduce dimensionality while preserving the most informative connectivity patterns for subsequent prediction.

3.2.4.1 Thresholded Estimator of ceFC

As shown in Chapter 2, the sample moment estimator of ceFC may fail to achieve consistent estimation in the high-dimensional setting, i.e., when the number of time points n is small relative to the number of brain regions p. To address this problem, we introduced a regularized estimator based on a thresholding operator that removes weak edge-pair interactions according to their magnitudes. Thresholding methods have been widely used in neuroimaging analysis to enhance interpretability and highlight more distinct connectivity patterns (Stam, 2004; De Vico Fallani et al., 2017; Garrison et al., 2015; Nicolini et al., 2020). Specifically, for a given cutoff value $\tau > 0$, we apply the thresholding operator elementwise to the ceFC estimator such that any entry with value $\widehat{\Theta}_{jk,st}$ where $|\widehat{\Theta}_{jk,st}| < \tau$ is truncated to zero. In this way, only relatively strong edge-pair interactions are retained. As established in Chapter 2, the resulting thresholded estimator yields a consistent estimator of ceFC even under high-dimensional conditions. In the present chapter, this thresholding step also serves as an initial screening procedure, substantially reducing the number of edge-pair features before subsequent dimension reduction and predictive modeling.

For each subject, thresholding the ceFC matrix produces a sparse and symmetric edge-by-edge representation. We retain the upper triangular entries excluding the diagonal, vectorize them, and concatenate these vectors from all subjects. To construct a shared reduced feature space for downstream analysis, we keep only those edge-pair features that remain nonzero after thresholding for every subject. This yields the reduced subject-by-edge-pair matrix $\widehat{\boldsymbol{\Theta}}_\tau$.

For the ID1000 data analyzed here, we select the threshold level $\tau = 0.01$ for each subject-specific ceFC matrix. With this choice, the intersection of nonzero edge-pair features across all subjects contains 62,225 entries. Therefore, the reduced matrix $\widehat{\boldsymbol{\Theta}}_\tau$ has 62,225 columns, whereas the node-centric matrix $\widehat{\boldsymbol{R}}_{\text{full}}$ remains in its original full form. This step substantially reduces the dimensionality of the edge-centric representation while preserving a common set of relatively strong edge-pair connections for all subjects.

3.2.4.2 Principal Component Analysis

Although thresholding substantially reduces the dimension of the subject-by-edge-pair matrix by removing weak edge-pair interactions, the resulting feature space remains too large for direct use in predictive modeling. It is therefore necessary to further compress the connectivity representations into a lower-dimensional set of features. To this end, we apply principal component analysis (PCA), a widely used dimension reduction technique in statistics and neuroimaging (Pearson, 1901; Hotelling, 1933). In general, PCA reduces the dimensionality of a dataset through a sequence of orthogonal linear combinations of variables, referred to as principal components (PCs), which are ordered by their explained variance. The first PC captures the largest possible variation in the original data, and each subsequent PC explains the largest remaining variation subject to orthogonality to the preceding components.

In this study, we apply PCA separately to the full subject-by-node-pair matrix $\widehat{\boldsymbol{R}}_{\text{full}}$ and the reduced subject-by-edge-pair matrix $\widehat{\boldsymbol{\Theta}}_{\tau}$. The resulting PCs are used as candidate predictors in the phenotype prediction models in the following analysis. Following common practice in the neuroimaging literature (Smith et al., 2015; Marek et al., 2022), we extract the top 100 PCs from $\widehat{\boldsymbol{R}}_{\text{full}}$, which together explain more than 90% of the total variation in the node-centric connectivity features. For the reduced subject-by-edge-pair matrix $\widehat{\boldsymbol{\Theta}}_{\tau}$, we retain between 10 and 50 PCs depending on the comparison scenarios considered in Section 3.3. This PCA step provides a compact representation of the major modes of variation in both node-centric and edge-centric connectivity and makes the subsequent modeling procedure computationally feasible.

3.2.5 Model

In brain-wide association studies (BWAS), a central objective is to understand how variation in brain structure and function is associated with complex phenotypic traits (Smith et al., 2015; Marek et al., 2022; Liu et al., 2023; Kang et al., 2024). In this setting, functional connectivity is often treated as a quantitative representation of brain organization and used as a set of predictors for behavioral or phenotypic outcomes. Given brain and phenotype data, one common modeling strategy is partial least squares (PLS) (Wold, 1975; Wold et al., 1984). This multivariate statistical technique has been widely applied in neuroimaging to identify shared variation between brain activity and behavior (McIntosh et al., 1996; Helmer et al., 2024; Krishnan et al., 2011; Popovic et al., 2020). In addition to PLS, we also consider baseline linear regression models in order to provide a simple reference for comparison. Model performance is evaluated by comparing predicted and observed phenotypic values using correlation and mean squared error (MSE), which are commonly used measures in

the neuroimaging and BWAS literature (Marek et al., 2022; Smith et al., 2015; Kang et al., 2024; Mantwill et al., 2022; Bzdok et al., 2021).

In this chapter, we investigate whether ceFC provides additional predictive information beyond conventional nFC. Specifically, we examine whether incorporating ceFC-derived features, together with nFC-derived features, improves the prediction of phenotypic outcomes relative to models based on nFC alone. Because both nFC and ceFC representations are high-dimensional relative to the available sample size, we use the principal components extracted from $\widehat{\boldsymbol{R}}_{\text{full}}$ and $\widehat{\boldsymbol{\Theta}}_{\tau}$ as low-dimensional inputs for model fitting. In this way, the comparison focuses on whether the dominant modes of variation captured by ceFC contribute predictive value beyond that already represented by nFC.

Because the predictors used in our analysis are principal components, the covariates derived from the same connectivity representation are mutually orthogonal. This structure has an important implication for model fitting. In particular, when the predictor matrix is orthonormal, the PLS model reduces to a linear regression, which is essentially a principal component regression model since the covariates are all PCs. More specifically, we show that when the input covariates are orthonormal, PLS will stop after the first iteration and will become identical to linear regression. On the other hand, when the predictor matrix $\boldsymbol{X}$ is orthogonal but not orthonormal, that is, $\boldsymbol{X}^\top \boldsymbol{X} = \boldsymbol{K}$ where $\boldsymbol{K}$ is a diagonal matrix but not identity, the relationship between PLS and OLS can still be simplified. In this case, the PLS coefficients $\boldsymbol{B}_{PLS}$ differ from the ordinary least squares (OLS) coefficients $\boldsymbol{B}_{OLS}$ by a scaling factor of

$$\left(\boldsymbol{Y}^\top \boldsymbol{X}\boldsymbol{K}\boldsymbol{X}^\top \boldsymbol{Y}\right)^{-1}\left(\boldsymbol{X}^\top \boldsymbol{Y}\,\boldsymbol{Y}^\top \boldsymbol{X}\boldsymbol{K}\right).$$

Therefore, when $\boldsymbol{X}$ is orthonormal so that $\boldsymbol{K} = \boldsymbol{I}$, the PLS and OLS estimators coincide. This result helps clarify why, in our setting, PLS and linear regression are expected to behave slightly differently when the predictors are given by PCs. We provide the PLS algorithm for orthogonal predictors in Algorithm 2, and the detailed derivation is presented in Appendix B.1.

The PLS model is constructed by extracting latent components that maximize the covariance between the predictors and the responses (Wold, 1975; Wold et al., 1984). In our analysis, the predictors are principal components derived from the connectivity patterns described above, so PLS is fitted to these component scores rather than to the original high-dimensional features. This formulation is consistent with the orthogonal structure discussed above for comparison with linear regression on these PCs. In practice, the number of PLS components is selected by minimizing the validation error. Predictive performance is evaluated using correlation and mean

squared error, and the corresponding results are reported in Section 3.3.

Algorithm 2 PLS with orthogonal predictors.

Input: An orthogonal matrix $\boldsymbol{X}$ with $\boldsymbol{X}^\top\boldsymbol{X} = \boldsymbol{K}$ where $\boldsymbol{K}$ is a diagonal matrix and response $\boldsymbol{Y}$.

For PLS component $l = 1, \ldots, L$:

1. Let w, q be the left and right singular vectors of $\boldsymbol{X}^\top\boldsymbol{Y}$, respectively.
2. Obtain scores
$$z = \boldsymbol{X}w = \frac{\boldsymbol{X}\boldsymbol{X}^\top\boldsymbol{Y}}{\sqrt{\boldsymbol{Y}^\top\boldsymbol{X}\boldsymbol{X}^\top\boldsymbol{Y}}}.$$
3. Obtain loadings
$$p = \boldsymbol{X}^\top z = \frac{\boldsymbol{X}^\top\boldsymbol{X}\boldsymbol{X}^\top\boldsymbol{Y}}{\sqrt{\boldsymbol{Y}^\top\boldsymbol{X}\boldsymbol{X}^\top\boldsymbol{Y}}}.$$
4. Obtain the rotation matrix
$$\boldsymbol{R} = \boldsymbol{W}(\boldsymbol{P}^\top\boldsymbol{W})^{-1} = (\boldsymbol{Y}^\top\boldsymbol{X}\boldsymbol{K}\boldsymbol{X}^\top\boldsymbol{Y})^{-1}\boldsymbol{X}^\top\boldsymbol{Y}\sqrt{\boldsymbol{Y}^\top\boldsymbol{X}\boldsymbol{X}^\top\boldsymbol{Y}}.$$
5. Obtain the PLS coefficients
$$\boldsymbol{B}_{PLS} = \boldsymbol{R}\boldsymbol{T}^\top\boldsymbol{Y} = (\boldsymbol{Y}^\top\boldsymbol{X}\boldsymbol{K}\boldsymbol{X}^\top\boldsymbol{Y})^{-1}(\boldsymbol{X}^\top\boldsymbol{Y}\boldsymbol{Y}^\top\boldsymbol{X}\boldsymbol{K})\boldsymbol{B}_{OLS},$$
where $\boldsymbol{B}_{OLS} = (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{Y}$.
6. Deflate data as
$$\boldsymbol{X} = \boldsymbol{X} - zp^\top, \qquad \boldsymbol{Y} = \boldsymbol{Y} - zq^\top.$$

3.3 Results

In this section, we present extensive numerical studies based on the ID1000 data set (Snoek et al., 2020) to investigate whether ceFC contributes additional predictive information for a range of phenotypic variables. The central goal is to compare predictive performance between two feature constructions: one using only subject-by-node-pair information, and the other combining subject-by-node-pair information with subject-by-edge-pair information. More specifically, we evaluate two approaches for selecting PC predictors. The first approach uses the top k PCs obtained from $\widehat{\boldsymbol{R}}_{\text{full}}$. The second approach uses the top k PCs from $\widehat{\boldsymbol{R}}_{\text{full}}$ together with the top q PCs from $\widehat{\boldsymbol{\Theta}}_\tau$. Through this design, we examine whether the inclusion of ceFC-related features improves prediction beyond the information contained in standard node connectivity-based representations.

To evaluate predictive performance, we use several complementary metrics. First, we compute the mean squared error (MSE) for each of the m phenotypic traits on the test set. Specifically, for each phenotypic trait y_j, the MSE is defined as

$\text{MSE}_j = \frac{1}{n}\sum_{i=1}^{n}(y_{ij} - \widehat{y}_{ij})^2$, where y_{ij} and $\widehat{y}_{ij}$ denote the observed and predicted values for subject i and trait j, respectively. To summarize prediction error across all response variables, we use the total test error $\sum_{j=1}^{m}\text{MSE}_j$. In addition to error-based measures, we also consider association-based metrics that are commonly reported in brain studies (Marek et al., 2022; Smith et al., 2015; Kang et al., 2024; Mantwill et al., 2022; Bzdok et al., 2021). In particular, we report the mean Pearson correlation and the mean Spearman correlation across all phenotypic traits. Pearson correlation reflects the strength of the linear association between observed and predicted values, whereas Spearman correlation reflects the strength of their monotone relationship. Taken together, these metrics provide a more complete evaluation of predictive performance across phenotypic outcomes.

We randomly partition the data into training, validation, and test sets with 447, 200, and 200 subjects, respectively. For each split, we tune the numbers of PCs selected from the subject-by-node-pair matrix $\widehat{\boldsymbol{R}}_{\text{full}}$ and the thresholded subject-by-edge-pair matrix $\widehat{\boldsymbol{\Theta}}_{\tau}$ based on validation performance. Specifically, we fit the candidate models on the training set, evaluate them on the validation set using the metrics introduced above, and then choose the combination of nFC PCs and ceFC PCs that performs best on the validation data. The selected model is subsequently evaluated on the test set to obtain the final prediction results for that split. To improve the stability of the comparison and reduce sensitivity to any single random partition, we repeat this procedure over 100 random splits and summarize the results across all splits. For ease of presentation, in the following figures, we denote the PCs from $\widehat{\boldsymbol{R}}_{\text{full}}$ by nFC PCs and the PCs from $\widehat{\boldsymbol{\Theta}}_{\tau}$ by ceFC PCs.

3.3.1 Linear Regression Results

We begin by evaluating linear regression models fitted separately for each phenotypic trait. To assess the overall predictive performance across all outcomes, we summarize the total validation MSE, defined as the sum of the trait-specific MSEs in Figure B.1. Since model performance depends on the number of nFC PCs and ceFC PCs included as predictors, we further examine the validation MSE over the full set of candidate combinations in Figure B.1a. As shown in Figure B.1a, lower validation MSE is concentrated in a relatively limited region of the tuning space, suggesting that prediction performance is sensitive to the choice of nFC and ceFC PCs.

To better understand this pattern, we next examine the validation MSE for several fixed numbers of ceFC PCs, as shown in Figure B.1b. Specifically, we compare models using nFC PCs alone with models that additionally include 1, 2, or 3 ceFC PCs. Across all four curves, the validation MSE gradually increases as the number of

nFC PCs becomes larger, indicating that including more nFC PCs does not improve predictive performance in this setting and may introduce additional noise. Moreover, for a fixed number of nFC PCs, models augmented with ceFC PCs generally yield slightly higher validation MSE than the model using nFC PCs alone. This pattern suggests that, under linear regression, the additional ceFC PCs considered here do not provide a clear predictive benefit and may slightly degrade model performance.

Based on the validation results, we select 5 nFC PCs and consider models with up to 1 ceFC PC for the final comparison. We then evaluate their generalization performance on the test set across 100 random splits, as summarized in Figure 3.1. The average test MSE shows a very small decrease when one ceFC PC is added to the model with 5 nFC PCs. However, the magnitude of this improvement is modest relative to the test MSE values. Overall, these results provide little evidence that adding a small number of ceFC PCs leads to a meaningful benefit in out-of-sample prediction under the linear regression model.

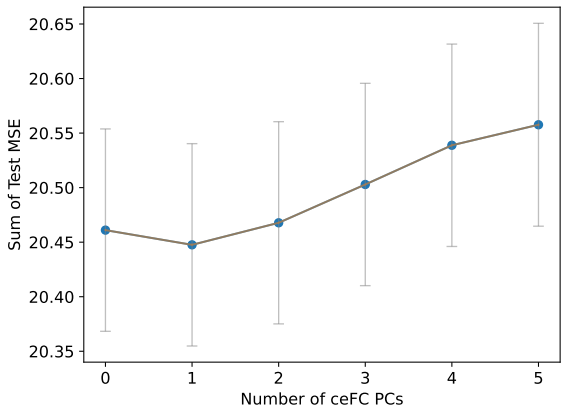


Figure 3.1: Sum of test MSEs using 5 nFC PCs and varying numbers of ceFC PCs. The gray bars depict standard errors across 100 random splits.

A more detailed breakdown by phenotypic trait is given in Figure 3.2, which compares three models: (1) 5 nFC PCs only, (2) 5 nFC PCs plus 1 ceFC PC, and (3) a baseline model that uses the training-set response mean. For most traits, adding 1 ceFC PC yields test MSE values that are very similar to those obtained using nFC PCs alone. A small reduction in test error is observed for several traits, including age and some Intelligence Structure Test measures, but the magnitude of the reduction is limited. Moreover, for a number of traits, the baseline mean-only model achieves the lowest test MSE. This pattern suggests that, for these phenotypes, neither node-based nor edge-based connectivity features provide substantial predictive information within the current linear regression framework.

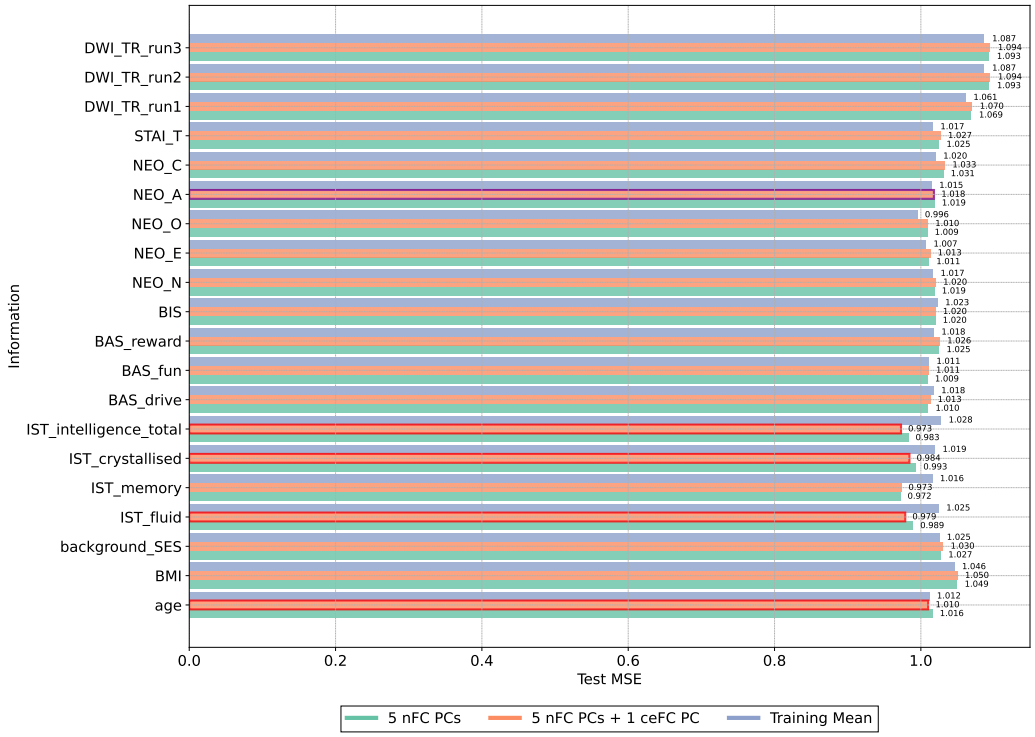


Figure 3.2: Test MSE for each phenotypic trait. The green, orange, and blue bars represent results using 5 nFC PCs only, 5 nFC PCs with 1 ceFC PC, and training response mean only, respectively. The highlighted red bar denotes the case in which adding the ceFC PC yields the smallest test MSE. The highlighted purple bar denotes the case in which only using the training response mean is the best, even though adding ceFC PCs yields a smaller test error in comparison to using nFC PCs only.

We next evaluate model performance using correlation-based measures. Since the Pearson and Spearman results show highly similar patterns, we present the Pearson correlation results in the main text and defer the Spearman correlation results to Appendix B.4. Figure B.2a displays the average validation Pearson correlation across all phenotypic traits for the full set of nFC and ceFC PC combinations (Figure B.5a for Spearman correlation). Both results indicate that a higher correlation is achieved when the model includes a moderate-to-large number of nFC PCs together with only a small number of ceFC PCs.

To examine this pattern more closely, Figure B.2b presents the validation Pearson correlation for models with 5 to 100 nFC PCs and 0 to 3 ceFC PCs (Figure B.5b for Spearman correlation). The correlation increases substantially as the number of nFC PCs grows from small to moderate values and then levels off. In contrast, adding ceFC PCs does not improve performance and generally leads to slightly lower Pearson correlation for a fixed number of nFC PCs. The Spearman correlation results, reported in Figure B.5b, show a similar qualitative pattern.

For the final evaluation on the test set, we fix the number of nFC PCs at the

value that performed well on the validation set (40 for Pearson correlation and 55 for Spearman correlation). We then consider models with 0 to 3 additional ceFC PCs. The resulting average test Pearson correlations are summarized in Figure 3.3. The mean correlation changes only slightly across different numbers of ceFC PCs. Overall, these results indicate that including ceFC PCs does not lead to a clear improvement in out-of-sample correlation under the linear regression model. The corresponding test Spearman correlation results, reported in Figure B.6, show a similar qualitative pattern. Together with the MSE-based analysis, these findings suggest that the additional ceFC PCs considered here do not provide meaningful predictive benefit within the linear regression framework.

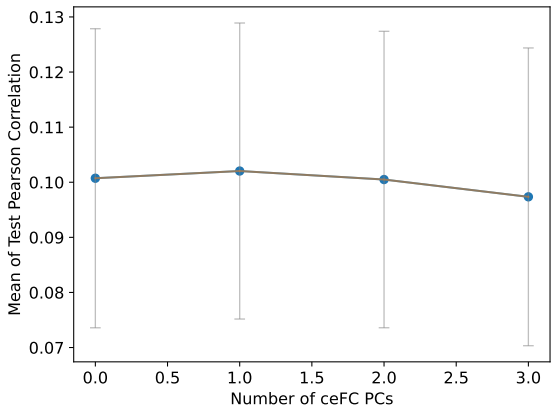


Figure 3.3: Mean test Pearson correlation using 40 nFC PCs and varying numbers of ceFC PCs. The gray bars depict standard errors across 100 random splits.

A more detailed trait-specific comparison is given in Figure 3.4, which contrasts the test Pearson correlation for each phenotype between the model using 40 nFC PCs only and the model using 40 nFC PCs plus 1 ceFC PC. For several traits, the model with 1 additional ceFC PC achieves a slightly higher correlation, and these cases are highlighted in red. However, the differences are generally small, and the direction of the change is not uniform across traits. In particular, modest improvements are observed for some intelligence-related measures and a few personality traits, whereas other phenotypes show little change or slight decreases in correlation. A similar pattern is also observed in Figure B.7 for Spearman correlation. Taken together, these trait-level results are consistent with the aggregate test-set analysis and with the earlier MSE-based findings, suggesting that the additional ceFC PCs considered here do not provide substantial predictive benefit beyond that captured by nFC PCs alone in linear regression.

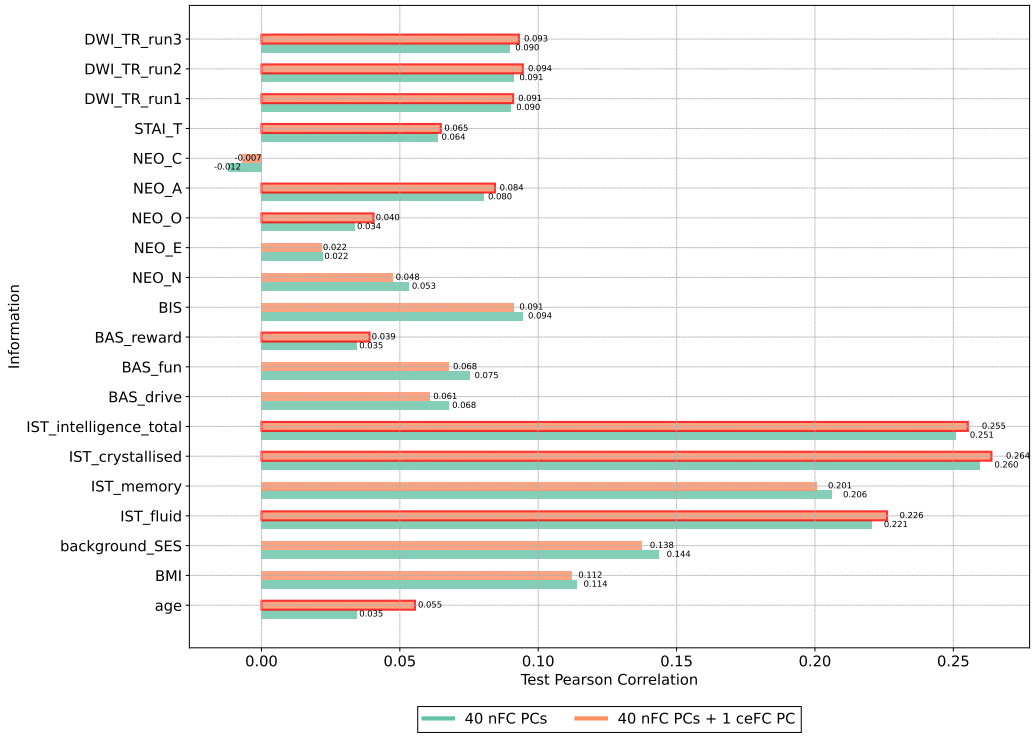


Figure 3.4: Test Pearson correlation for each phenotypic trait. The green and orange bars represent results using 40 nFC PCs only, and 40 nFC PCs with 1 ceFC PC, respectively. The highlighted red bar denotes the case in which adding the ceFC PC yields a larger test correlation.

3.3.2 PLS Results

We next apply the same analysis pipeline using partial least squares (PLS) regression, with the combined set of nFC and ceFC PCs as predictors. Figure B.3a summarizes the total validation MSE across all phenotypic traits over a range of PLS components and predictor combinations. Among the candidate models considered here, the lowest validation MSE is attained when the number of PLS components is set to 2. We therefore focus on the 2-component PLS model and examine the validation MSE over different combinations of nFC and ceFC PCs in Figure B.3b. The figure shows that adding ceFC PCs can reduce the validation MSE when the number of nFC PCs is relatively small. However, once the number of nFC PCs reaches the range that yields the best overall performance, approximately 40 to 60 PCs, the inclusion of ceFC PCs provides little additional benefit and in some cases leads to slightly larger validation error.

To assess whether this pattern persists on the test set, we fix the model at 2 PLS components and 45 nFC PCs, and then vary the number of ceFC PCs, as shown in Figure 3.5. The average test MSE changes only modestly across the candidate

numbers of ceFC PCs. Although the model including 1 ceFC PC gives a slightly lower average test MSE than the model using nFC PCs alone, the improvement is small relative to the test MSE values. Overall, the test-set results provide limited evidence that adding ceFC PCs yields a meaningful gain in predictive performance under PLS models.

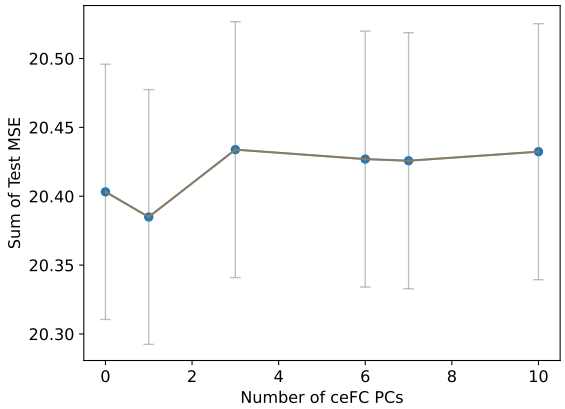


Figure 3.5: Sum of test MSEs over all phenotypic traits with standard errors in gray bars across 100 random splits based on 2 PLS components, 45 nFC PCs and varying numbers of ceFC PCs.

A more detailed trait-specific comparison is provided in Figure 3.6, which contrasts the test MSE for each phenotype across the selected PLS models and the baseline model that uses the training-set response mean. Similar to the linear regression results, the inclusion of ceFC PCs does not lead to a consistent reduction in test error across traits. Modest improvements are observed for a small number of intelligence- and personality-related measures, but these gains are limited in magnitude. For many phenotypes, the mean-only baseline remains competitive and in some cases achieves lower test error than the connectivity-based models. Taken together, these findings suggest that, within the current PLS framework, the additional ceFC PCs considered here offer only limited predictive benefit beyond that already captured by nFC PCs.

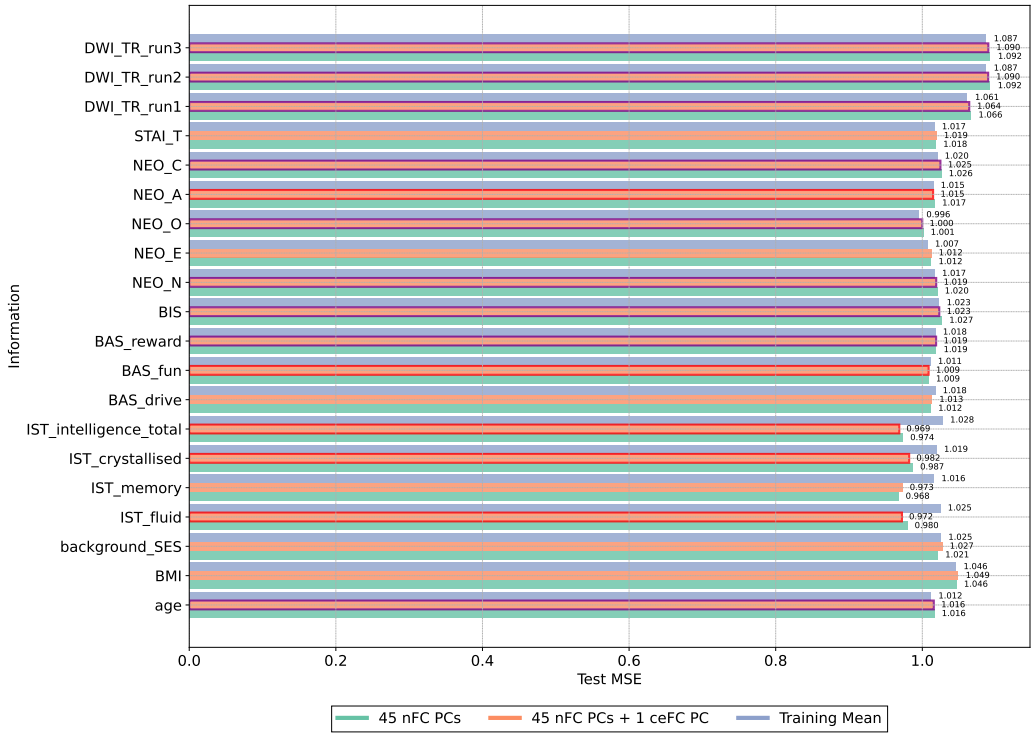


Figure 3.6: Test MSE for each phenotypic trait. The green, orange, and blue bars represent results using 45 nFC PCs only, 45 nFC PCs with 1 ceFC PC, and training response mean only, respectively. The highlighted red bar denotes the case in which adding the ceFC PC yields the smallest test MSE. The highlighted purple bar denotes the case in which only using the training response mean is the best, even though adding ceFC PC yields a smaller test error in comparison to using nFC PCs only.

We next evaluate the PLS models using correlation-based measures. Since the Pearson and Spearman results show highly similar patterns, we present the Pearson correlation results in the main text and the Spearman correlation results in Appendix B.5. Figure B.4a summarizes the mean validation Pearson correlation across all phenotypic traits over a range of PLS components and predictor combinations. Among the candidate models considered here, the highest validation correlation is achieved when the number of PLS components is set to 5. We therefore focus on the 5-component PLS model and examine the validation Pearson correlation over different combinations of nFC and ceFC PCs in Figure B.4b. The correlation increases substantially as the number of nFC PCs grows from small to moderate values and then gradually levels off. In contrast, adding ceFC PCs does not improve validation performance and generally leads to slightly lower correlation for a fixed number of nFC PCs. The corresponding Spearman correlation results, reported in Figure B.8, show the same qualitative pattern.

We then assess the selected PLS model on the test set. Fixing the number of PLS components at 5 and varying the number of ceFC PCs, Figure 3.7 shows that

the mean test Pearson correlation remains nearly flat and tends to decrease slightly as additional ceFC PCs are included. The associated standard errors overlap substantially across models, indicating that these differences are small. A more detailed trait-specific comparison is provided in Figure 3.8, which contrasts the test Pearson correlation for each phenotype between the model using nFC PCs only and the model including an additional ceFC PC. For several traits, the model with the additional ceFC PCs yields slightly higher correlation, and these cases are highlighted in red. However, the improvement is generally modest, and many traits show little change or slight decreases. Overall, these results are consistent with the MSE-based analysis and suggest that, within the current PLS framework, the additional ceFC PCs considered here do not provide substantial predictive benefit beyond that captured by nFC PCs alone. The corresponding test Spearman correlation results in Figure B.9 and Figure B.10 indicate the same qualitative conclusion.

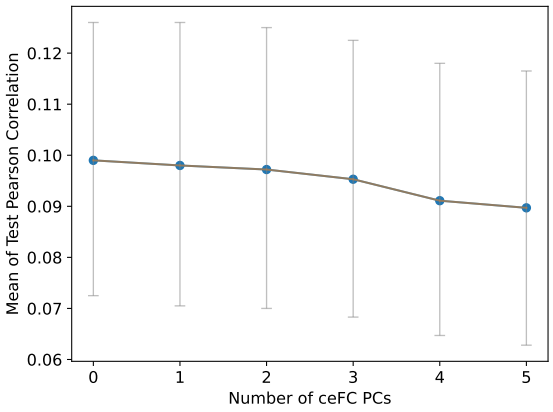


Figure 3.7: Mean test Pearson correlation using 100 nFC PCs and varying numbers of ceFC PCs under 5 PLS components. The gray bars depict standard errors across 100 random splits.

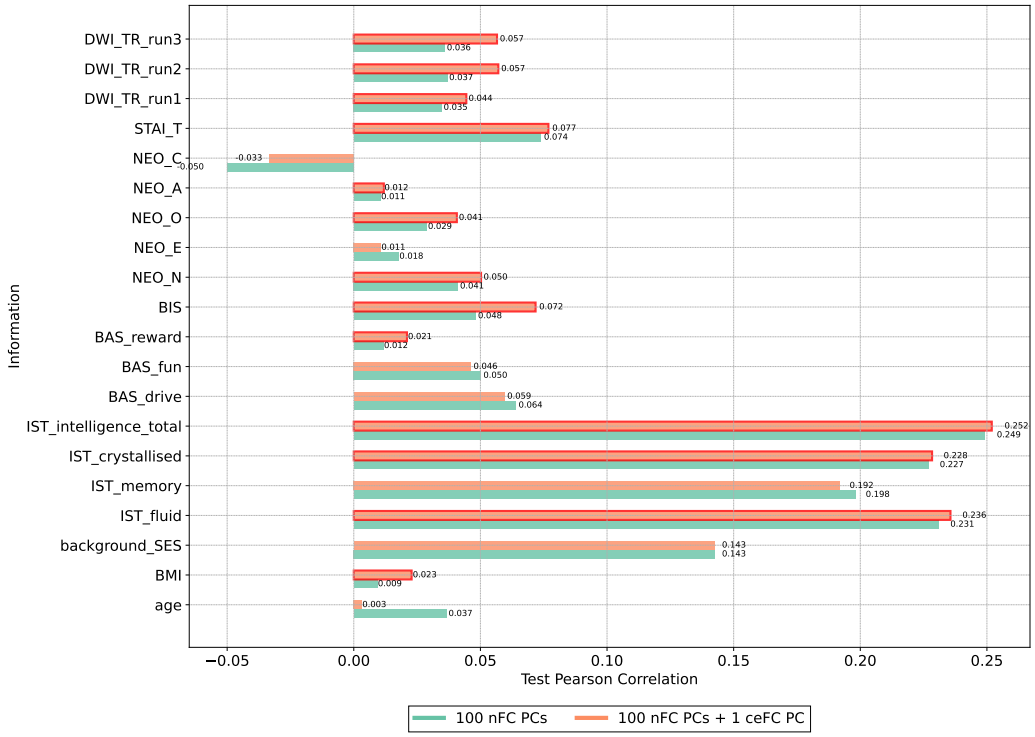


Figure 3.8: Test Pearson correlation for each phenotypic trait with 5 PLS components. The green and orange bars represent results using 100 nFC PCs only, and 100 nFC PCs with 1 ceFC PC, respectively. The highlighted red bar denotes the case in which adding the ceFC PCs yields a larger test correlation.

3.3.3 Explanation

The consistently limited improvement observed after augmenting nFC predictors with ceFC predictors suggests that ceFC provides little additional information for the prediction of the phenotypic traits considered in this study. This observation motivates a closer examination of the relationship between the two connectivity methods. In this subsection, we explore a possible explanation for this finding by focusing on the intrinsic statistical dependence between nFC and ceFC.

As discussed in Chapter 2, ceFC is not an independent measurement, but is mathematically constructed from the entries of the nFC matrix. This construction implies a strong structural relationship between the two representations and suggests that the information carried by ceFC may overlap substantially with that already encoded in nFC. If this is the case, then the principal components (PCs) extracted from the two matrices should also exhibit strong dependence.

To investigate this redundancy more directly, we conduct a simulation study based on a reduced set of brain nodes. Specifically, we randomly select 10 nodes from the full set of 200 nodes and use the corresponding sample correlation matrix from the

fMRI time series as the subject-specific ground-truth nFC matrix. For each sample size $N \in \{10, 20, 40, 60\}$, we then construct the subject-by-node-pair matrix $\widehat{\boldsymbol{R}}_{\text{full}}$ and extract its nFC PCs. Under the assumption of multivariate normality, we compute the corresponding ground-truth ceFC for each subject using the established second- and fourth-order moment formulas between edge jk and edge st as $\Theta^*_{jk,st} = \sigma_{js}\sigma_{kt} + \sigma_{jt}\sigma_{ks}$ by Isserlis's theorem (Isserlis, 1918), where σ_{jk} denotes the covariance between node j and node k. We assemble the subject-by-edge-pair matrix $\widehat{\boldsymbol{\Theta}}_{\text{full}}$ without thresholding, since the reduced node set keeps the computation manageable, and extract its leading PCs. For each ceFC PC, we regress the true ceFC PC on the full set of nFC PCs and compute the coefficient of determination, R^2.

Figure 3.9 summarizes the above results. Across all sample sizes, the R^2 values are uniformly high for most ceFC PCs, indicating that a large proportion of the variation in ceFC can be explained by linear combinations of nFC PCs. Although the explanatory power decreases slightly for some higher-order ceFC PCs as the sample size becomes larger, the overall dependence remains strong. This pattern supports that ceFC is largely redundant with nFC at the level of principal component representations. It also provides a plausible explanation for the prediction results reported above. When a model already includes a sufficiently rich set of nFC PCs, adding ceFC PCs contributes little new information and therefore yields little improvement in the prediction of phenotypic traits.

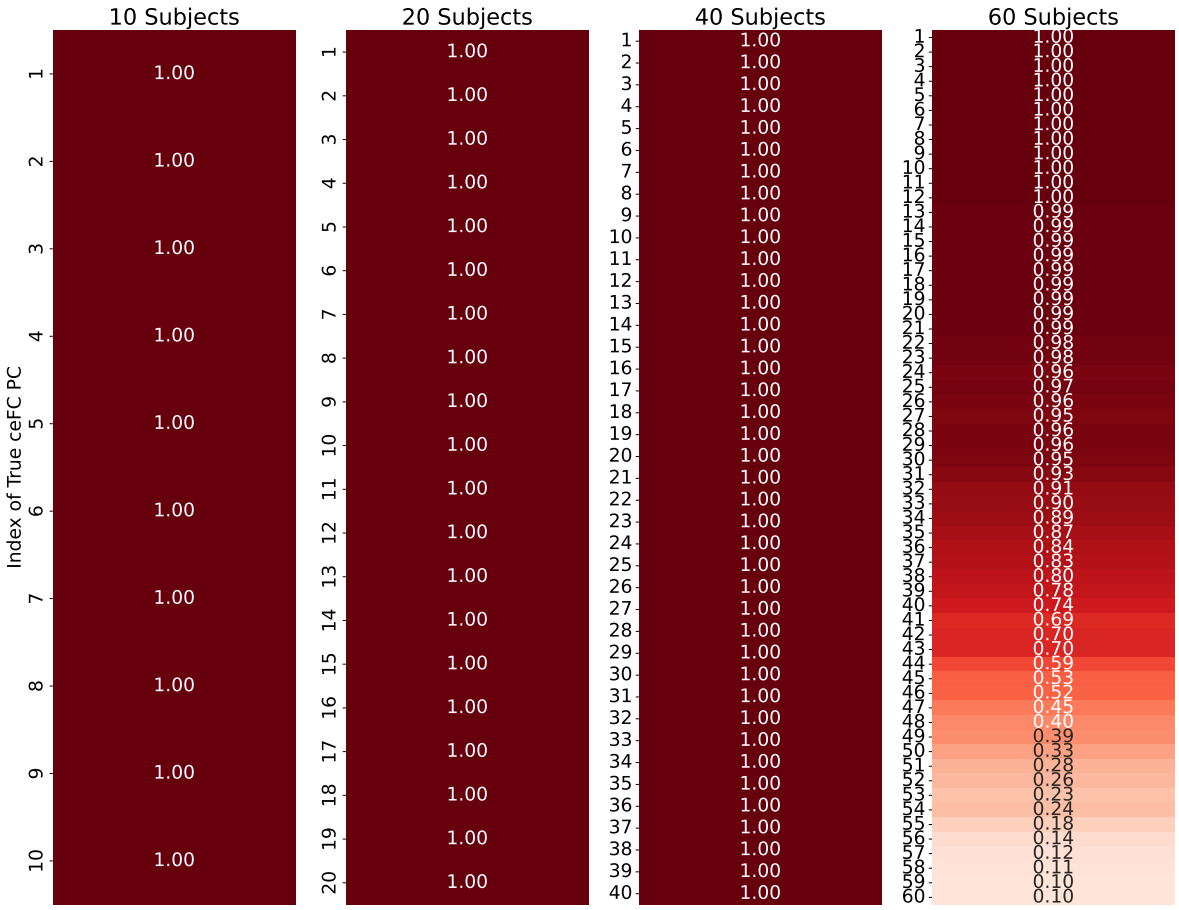


Figure 3.9: R^2 from the linear regression of each true ceFC PC on all nFC PCs with 10 nodes.

To examine whether the same redundancy persists in the full setting, we extend the analysis to all 200 brain nodes. Following the same general strategy as above, we construct the subject-by-edge-pair matrix $\widehat{\mathbf{\Theta}}_\tau$ using the regularized estimator of ceFC and extract its leading PCs. We then regress each ceFC PC on the full set of nFC PCs and record the resulting coefficient of determination, R^2. Figure 3.10 summarizes these results. The R^2 values remain high across the leading ceFC components, showing that much of the dominant ceFC variation can be explained by the nFC representation through a linear relationship. Moreover, the top 15 ceFC PCs account for nearly 90% of the total variation in $\widehat{\mathbf{\Theta}}_\tau$, so this dependence is not limited to a negligible part of the ceFC space, but instead characterizes most of its major variation. Taken together with the reduced-node analysis above, these results indicate that ceFC is highly redundant with nFC at the level of principal component representations. This redundancy offers a natural explanation for why adding ceFC PCs leads to, at most, limited improvement in phenotypic prediction when sufficient nFC PCs are already included in the model.

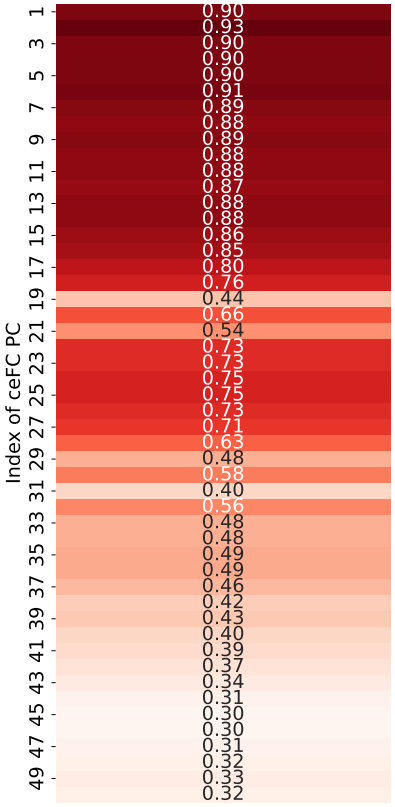


Figure 3.10: R^2 from the regressing each ceFC PC, extracted from the ceFC estimate, on all nFC PCs using the full set of 200 brain nodes.

CHAPTER 4

Spatio-temporally Varying Coefficient Model via Deep Neural Networks with Applications in fMRI

4.1 Introduction

In Chapters 2 and 3, we focused on connectivity patterns among brain networks, with the goal of gaining a new perspective on brain organization. Beyond connectivity analysis, another central objective in neuroimaging studies is to characterize the associations between brain imaging responses and external factors within anatomical space. This is scientifically important because such associations can localize covariate effects over the brain image domain and further provide interpretable markers for neurological disorders and functional brain architecture (Smith et al., 2015; Woo et al., 2017; Smith et al., 2009). At the same time, fMRI data often produce high-dimensional observations with complex dependencies over space and time, imposing challenges for regression tasks.

To address these challenges, various regularization techniques and structured models have been proposed. Among them, the spatially varying coefficient model (SVCM) has emerged as a powerful framework. SVCMs represent covariate effects as coefficient functions varying over the image domain in image responses, thereby capturing localized effects and accommodating spatial smoothness, sparsity, jump discontinuities, and heterogeneity (Zhu et al., 2014; Li et al., 2017, 2021). These methods have been widely applied in neuroimaging studies to uncover the spatial patterns of covariate effects within brain networks.

Building on this literature, Zhang et al. (2023) proposed image response regression via neural networks (IRRNN), which applies deep neural networks to approximate spatially varying coefficient functions. This approach integrates spatial smoothness and accounts for subject heterogeneity while preserving the interpretation of the estimated covariate effects. Compared with local smoothing or spline-based methods, it provides a more flexible way to capture complex spatial patterns. Moreover, by learning the coefficient functions from spatial coordinates as inputs, IRRNN (Zhang

et al., 2023) alleviates the high-dimensionality challenge in traditional image regression, such that the effective sample size depends on the number of voxels rather than only on the number of subjects, which are often limited in real fMRI datasets. This feature makes the neural-network-based estimation more feasible in neuroimaging applications.

Despite this progress, the existing methods primarily model coefficient functions over a spatial domain and do not explicitly account for temporal evolution. However, dynamic fMRI data are inherently indexed by both space and time. Recent neuroimaging studies increasingly involve continuous fMRI scans in which brain activity is recorded during naturalistic paradigms, such as movie watching. These stimuli are dynamic, multimodal, and context-rich, enabling a deeper understanding of brain function under ecologically valid conditions that closely resemble real-world situations (Sonkusare et al., 2019; Finn and Bandettini, 2021; Matusz et al., 2019). Prior studies have shown that naturalistic stimuli engage widespread brain networks, revealing synchronized activity across individuals (Hasson et al., 2004), hierarchical processing over temporal receptive windows (Hasson et al., 2008; Honey et al., 2012; Lerner et al., 2011), and large-scale network reconfiguration during narrative comprehension (Simony et al., 2016), thereby facilitating higher-order functions such as social cognition and perception (Lahnakoski et al., 2012; Baldassano et al., 2017). Because naturalistic covariates, such as visual, audio, and emotional features, are scene-dependent, ignoring temporal evolution may overlook critical structures in the underlying effects. In this way, the effect of a scene-related covariate may vary across scenes, interact with spatial patterns, and exhibit a non-separable spatio-temporal structure.

Motivated by these considerations, this chapter develops a spatio-temporally varying coefficient model (STVCM) for dynamic image response data. The main idea is to model each regression coefficient as a function of both spatial location and temporal input, so that the association between covariates and image responses can evolve under dynamically changing stimuli. This formulation is particularly relevant for naturalistic fMRI data, where scene-related effects can vary not only across brain regions but also over time. Since the interaction between spatial patterns and temporal trajectories is often non-linear and non-separable, traditional approaches, such as spatio-temporal splines or methods based on Gaussian processes, may face challenges in both computational scalability and their ability to capture such complex interactions. Inspired by the use of deep neural networks in IRRNN (Zhang et al., 2023) to address these challenges in the spatial domain, we extend this idea to the spatio-temporal setting and propose spatio-temporal image response regression via neural networks (ST-IRRNN) for estimating the coefficient functions in the STVCM. Since the true form of spatio-temporal dependence in real data, such as fMRI scans, often

remains unknown, we develop several ST-IRRNN variants that combine spatial and temporal information in different architectural forms. The additive and multiplicative ST-IRRNN variants are intended for settings in which the spatio-temporal interaction is approximately separable, while the concatenative architecture in ST-IRRNN can directly learn from the joint spatio-temporal input and is therefore better suited to more general non-separable patterns.

Beyond these feed-forward neural network applications, recent developments in transformer architectures have provided another promising direction for modeling complex dependence in sequential data. Built on the self-attention mechanism, transformers have achieved strong performance in sequence modeling (Vaswani et al., 2017; Wolf et al., 2020), effectively capturing long-range dependencies and flexibly aggregating information across positions in text (Devlin et al., 2019; Radford et al., 2021; Yang et al., 2019) and image data (Dosovitskiy et al., 2021; Liang et al., 2021; Zhou et al., 2018). Transformer-based models have also been increasingly adopted in neuroimaging studies for spatio-temporal representation learning (Asadi et al., 2023; Kim et al., 2023; Gao et al., 2026) and for disorder identification using fMRI data (Deng et al., 2022; Niu et al., 2020; Dai et al., 2024).

In light of these developments, we further introduce a transformer-based ST-IRRNN formulation in this chapter. To enrich the temporal input before it is integrated with spatial information, we apply Fourier features (Tancik et al., 2020) as a temporal representation, helping the model encode nonlinear and potentially high-frequency temporal variation that may be difficult to learn efficiently from a raw scalar time coordinate alone (Tancik et al., 2020; Rahaman et al., 2019; Luo et al., 2022). The resulting Fourier-enhanced temporal representation is then combined with a spatial projection to form a joint embedding, which is processed by a transformer encoder (Vaswani et al., 2017). In contrast to the additive and multiplicative variants, this architecture does not impose a prespecified separable spatio-temporal structure. Additionally, unlike the concatenative variant, it uses self-attention to capture more structured temporal dependence, thus providing a more flexible framework for learning complex dependence between space and time directly from the data.

One advantage of the proposed ST-IRRNN approaches is that they remain interpretable at the coefficient-function level, which facilitates insight into evolving activation patterns across regions and over time. Meanwhile, these variants are flexible to accommodate spatio-temporally varying effects, including approximately separable trends as well as genuinely non-separable interactions. The proposed methods are evaluated under simulation settings motivated by Zhang et al. (2023) and extended to the spatio-temporal domain. These settings are designed to represent qualitatively different forms of spatio-temporal dependence, including additive, multiplicative, and

more dynamic non-separable patterns. These studies show that a matching structured ST-IRRNN variant tends to capture the true dynamically varying effects with lower error and more robust performance. Even when the spatio-temporal effects are non-separable, the proposed ST-IRRNN approaches can still improve over fixed-window IRRNN baselines in both estimation accuracy and temporal stability. We further demonstrate the practical value of the proposed framework through an application to a real fMRI dataset collected during movie watching (Chen et al., 2017). In this study, we show that the estimated spatio-temporally varying main effects associated with audio, visual, and emotional stimuli reveal dynamic scene-dependent patterns across brain systems. These fitted coefficient maps provide scientifically meaningful brain images showing how different stimulus features modulate brain activity over time, and illustrate the potential of the proposed framework for uncovering dynamic brain–covariate associations under continuous stimulation.

4.2 Model and Estimation

In this section, we formulate an STVCM for image response regression and introduce an estimation procedure that integrates flexible neural network approximators to estimate each key component within the model. While inspired by Zhang et al. (2023), our method is designed for dynamic imaging data over a joint spatio-temporal domain.

4.2.1 Image Response Regression with Spatio-Temporally Varying Coefficients

Let $\mathcal{S} \subset \mathbb{R}^D$ denote a D-dimensional spatial domain observed on a grid of V locations $\{\boldsymbol{s}_v\}_{v=1}^V$ and $\mathcal{T} \subset \mathbb{R}$ denote the temporal domain observed over time $t \in [0,1]$. We consider imaging data collected from N subjects. For each subject $i = 1, \ldots, N$, we observe the scalar image measurement $y_i(\boldsymbol{s}_v, t) \in \mathbb{R}$ at the spatial location $\boldsymbol{s}_v, v = 1, \ldots, V$ and time t. Let $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{iJ})^\top \in \mathbb{R}^J$ represent a J-dimensional covariate vector at time t associated with the i-th subject.

To characterize the dynamic heterogeneity of covariate effects across both space and time, we propose the following spatio-temporally varying coefficient model (STVCM):

$$y_i(\boldsymbol{s}_v, t) = \sum_{j=1}^{J} x_{ij}(t)\beta_j(\boldsymbol{s}_v, t) + \alpha_i(\boldsymbol{s}_v, t) + \epsilon_i(\boldsymbol{s}_v, t), \quad \mathrm{Var}\{\epsilon_i(\boldsymbol{s}_v, t)\} = \sigma^2(\boldsymbol{s}_v, t) \quad (4.2.1)$$

Here, $\beta_j(\boldsymbol{s}_v, t)$ is the varying coefficient function representing the population-level main effect of covariate x_j at location $\boldsymbol{s}_v$ and time t. $\alpha_i(\boldsymbol{s}_v, t)$ denotes the subject-specific deviation, which captures the intrinsic spatio-temporal variability of subject i that is not explained by the covariates. The error term $\epsilon_i(\boldsymbol{s}_v, t)$ is assumed to be an independent random error with mean zero and a spatio-temporally varying variance $\sigma^2(\boldsymbol{s}_v, t)$.

Unlike the SVCM, which treats observations as static images, the proposed STVCM in model (4.2.1) explicitly incorporates the temporal dimension and allows for the investigation of how covariate effects evolve over time. We assume that the functions $\beta_j(\boldsymbol{s}_v, t)$, $\alpha_i(\boldsymbol{s}_v, t)$, and $\sigma^2(\boldsymbol{s}_v, t)$ are piecewise smooth over the domain $\mathcal{S}\times\mathcal{T}$. In addition, the main effect function $\beta_j(\boldsymbol{s}_v, t)$ is assumed to be sparse, in the sense that its value is zero over certain regions at some time. Such a sparsity assumption is common in neuroimaging studies (Achard and Bullmore, 2007; Bullmore and Sporns, 2012; De Vico Fallani et al., 2017; Garrison et al., 2015; Nicolini et al., 2020; Li et al., 2021). Similarly, we impose a sparsity assumption on the individual effect $\alpha_i(\boldsymbol{s}_v, t)$, while allowing the sparse patterns to vary across subjects. That is, for each subject i, the zero-valued regions of $\alpha_i(\boldsymbol{s}_v, t)$ can differ across both space and time.

Next, we establish the identifiability of model (4.2.1).

Proposition 4.1. For any $t \in [0, 1]$ and any full rank matrix $\mathbf{X}(t) \in \mathbb{R}^{N\times J}$, let $\boldsymbol{\Lambda}$ denote the collection of all spatio-temporally varying coefficients $\boldsymbol{\lambda}(\mathbf{s}, t)$ in model (4.2.1), such that $\mathbf{X}(t)^T\boldsymbol{\alpha}(\mathbf{s}, t) = \mathbf{0}$ for all $\mathbf{s} \in \mathcal{S}$. Let $\pi\{\mathbf{y}(\mathbf{s}, t) \mid \mathbf{X}(t), \boldsymbol{\lambda}(\mathbf{s}, t)\}$ be the probability density function of $\mathbf{y}(\mathbf{s}, t)$ given $\mathbf{X}(t)$ and $\boldsymbol{\lambda}(\mathbf{s}, t)$ under model (4.2.1). For any $\widetilde{\boldsymbol{\lambda}}, \bar{\boldsymbol{\lambda}} \in \boldsymbol{\Lambda}$, if $\pi\{\mathbf{y}(\mathbf{s}, t) \mid \mathbf{X}(t), \widetilde{\boldsymbol{\lambda}}(\mathbf{s}, t)\} = \pi\{\mathbf{y}(\mathbf{s}, t) \mid \mathbf{X}(t), \bar{\boldsymbol{\lambda}}(\mathbf{s}, t)\}$ for all $\mathbf{s}$ and t, then $\widetilde{\boldsymbol{\lambda}} = \bar{\boldsymbol{\lambda}}$.

The proof is provided in Appendix C.1.

4.2.2 Varying Coefficients Approximation

Estimating each component in Model (4.2.1) is challenging due to the high dimensionality of the spatio-temporal domain. Traditional methods, such as kernel smoothing and spline regression, often suffer from the curse of dimensionality and become computationally inefficient when the spatio-temporal dimension is large (Cressie and Wikle, 2011; Zhu et al., 2014). To address this difficulty, Zhang et al. (2023) employed deep neural networks (DNNs) to model spatially varying coefficients in image regression by treating them as functions of spatial locations and exploiting the universal approximation property of DNNs. Their framework is referred to as image response regression via neural networks (IRRNN) (Zhang et al., 2023). Motivated by this idea, we adopt a similar strategy to model the spatio-temporal coefficient

functions while preserving the interpretability of the outputs as varying coefficients. Specifically, we approximate the unknown functions by neural networks evaluated at the spatio-temporal coordinate $(\boldsymbol{s}_v, t)$.

Motivated by the IRRNN framework (Zhang et al., 2023), we propose the spatio-temporal image response regression via neural networks (ST-IRRNN). To accommodate different forms of spatio-temporal dependence, we consider four ST-IRRNN architectures for approximating coefficient functions, namely, concatenative ST-IRRNN, additive ST-IRRNN, multiplicative ST-IRRNN, and transformer-based ST-IRRNN.

Concatenative ST-IRRNN We begin with a general concatenative spatio-temporal structure. Specifically, we consider a fully connected feed-forward neural network that takes the spatial and temporal coordinates jointly as input. Let $\boldsymbol{u} = (\boldsymbol{s}, t)^\top \in \mathbb{R}^{D+1}$ denote the combined spatio-temporal input. A neural network with L hidden layers can be written recursively as

$$\begin{aligned} \boldsymbol{h}^{(0)}(\boldsymbol{u}) &= \boldsymbol{u}, \\ \boldsymbol{h}^{(l)}(\boldsymbol{u}) &= \sigma_l\Big(\boldsymbol{W}_{l-1}\boldsymbol{h}^{(l-1)}(\boldsymbol{u}) + \boldsymbol{b}_{l-1}\Big), \quad l = 1, \ldots, L, \\ N(\boldsymbol{s}, t \mid \boldsymbol{\theta}) &= \boldsymbol{W}_L\boldsymbol{h}^{(L)}(\boldsymbol{u}) + \boldsymbol{b}_L, \end{aligned} \tag{4.2.2}$$

where $\boldsymbol{\theta} = \{\boldsymbol{W}_l, \boldsymbol{b}_l\}_{l=0}^{L}$ represents the collection of weight and bias parameters, and $\sigma_l(\cdot)$ is an elementwise nonlinear activation function, such as the Leaky ReLU or the sigmoid function. Under this formulation, the spatial coordinate $\boldsymbol{s}$ and the temporal input t are directly concatenated and passed through a single neural network. This architecture allows the model to learn general non-separable spatio-temporal dependence without imposing an explicit separability assumption.

Additive ST-IRRNN When an additive spatio-temporal structure is appropriate, we consider the following separable neural network

$$N(\boldsymbol{s}, t \mid \boldsymbol{\theta}) = N_{\boldsymbol{s}}(\boldsymbol{s} \mid \boldsymbol{\theta}_{\boldsymbol{s}}) + N_t(t \mid \boldsymbol{\theta}_t). \tag{4.2.3}$$

Here, $N_{\boldsymbol{s}}(\boldsymbol{s} \mid \boldsymbol{\theta}_{\boldsymbol{s}})$ denotes a neural network that takes the spatial coordinate $\boldsymbol{s}$ as input and models the spatial component. Similarly, $N_t(t \mid \boldsymbol{\theta}_t)$ denotes a neural network that takes time t as input and models the temporal component. Under this formulation, the spatial and temporal effects are represented by two networks and are combined additively at the output layer. This structure allows the spatial and temporal components to contribute separately, without explicitly modeling their interaction, and

is therefore suitable for settings in which the spatio-temporal dependence is approximately additive.

Multiplicative ST-IRRNN When a multiplicative spatio-temporal structure is appropriate, we consider the following separable neural network

$$N(\boldsymbol{s}, t \mid \boldsymbol{\theta}) = N_{\boldsymbol{s}}(\boldsymbol{s} \mid \boldsymbol{\theta}_{\boldsymbol{s}}) \times N_t(t \mid \boldsymbol{\theta}_t). \tag{4.2.4}$$

Here, $N_{\boldsymbol{s}}(\boldsymbol{s} \mid \boldsymbol{\theta}_{\boldsymbol{s}})$ and $N_t(t \mid \boldsymbol{\theta}_t)$ denote neural networks that take the spatial coordinate $\boldsymbol{s}$ and temporal coordinate t as input, respectively. Under this formulation, the spatial and temporal effects are represented by two networks and are combined multiplicatively at the output layer. This structure allows the temporal effect to modulate the spatial effect over time and is therefore suitable for settings in which the spatio-temporal dependence is approximately separable but still exhibits interaction between the two components.

Transformer-based ST-IRRNN To further enhance flexibility in modeling complex spatio-temporal dependence, we also consider a transformer-based architecture for $N(\boldsymbol{s}, t \mid \boldsymbol{\theta})$, inspired by the literature that self-attention provides an effective mechanism for modeling structured dependence in sequential data and for capturing long-range interactions that may be difficult to represent with standard feed-forward architectures (Vaswani et al., 2017; Wolf et al., 2020). Accordingly, unlike the additive, multiplicative, and concatenative variants above, this architecture combines learned spatial and temporal representations with a transformer encoder to model dependence in the resulting joint representation (Vaswani et al., 2017).

Specifically, let $N_{\boldsymbol{s}}(\boldsymbol{s} \mid \boldsymbol{\theta}_{\boldsymbol{s}}) \in \mathbb{R}^{d_s}$ denote a spatial projection network that maps the spatial coordinate $\boldsymbol{s} \in \mathbb{R}^D$ to a d_s-dimensional representation. To enrich the temporal input before integrating it with the spatial component, we represent the scalar time coordinate t using Fourier features:

$$\boldsymbol{\phi}(t) = \Big[t, \sin(\omega_1 t), \cos(\omega_1 t), \ldots, \sin(\omega_K t), \cos(\omega_K t)\Big]^\top \in \mathbb{R}^{1+2K},$$

where $\omega_k = 2\pi k$ for $k = 1, \ldots, K$ are the frequency parameters (Tancik et al., 2020). This Fourier feature mapping enriches the temporal representation and helps the model encode nonlinear and potentially high-frequency temporal variation that may be difficult to learn efficiently from a raw scalar time coordinate alone (Tancik et al., 2020; Rahaman et al., 2019; Luo et al., 2022). We then define a temporal projection network

$$N_t\Big(\boldsymbol{\phi}(t) \mid \boldsymbol{\theta}_t\Big) \in \mathbb{R}^{d_t}.$$

The resulting spatial and temporal representations are concatenated to form a joint embedding,

$$\boldsymbol{z}(\boldsymbol{s},t) = \mathrm{Concat}\Big(N_{\boldsymbol{s}}(\boldsymbol{s} \mid \boldsymbol{\theta}_{\boldsymbol{s}}), N_t\Big(\boldsymbol{\phi}(t) \mid \boldsymbol{\theta}_t\Big)\Big) \in \mathbb{R}^{d_s+d_t}.$$

This joint embedding is then projected into the transformer hidden space as

$$\boldsymbol{h}^{(0)}(\boldsymbol{s},t) = \boldsymbol{W}_{\mathrm{in}}\boldsymbol{z}(\boldsymbol{s},t) + \boldsymbol{b}_{\mathrm{in}} \in \mathbb{R}^{d_{\mathrm{model}}},$$

where d_{model} denotes the hidden dimension of the transformer.

The projected representations are subsequently passed through a transformer encoder with L layers. Let $\boldsymbol{H}^{(l-1)}$ denote the hidden representation at layer $l-1$. The multi-head self-attention (MHA) mechanism at the lth layer is defined as

$$\mathrm{MHA}\Big(\boldsymbol{H}^{(l-1)}\Big) = \mathrm{Concat}(\mathrm{head}_1, \ldots, \mathrm{head}_H)\boldsymbol{W}^O,$$

where each attention head is computed by scaled dot-product attention,

$$\mathrm{head}_h = \mathrm{softmax}\left(\frac{\boldsymbol{Q}_h\boldsymbol{K}_h^\top}{\sqrt{d_k}}\right)\boldsymbol{V}_h,$$

with $\boldsymbol{Q}_h = \boldsymbol{H}^{(l-1)}\boldsymbol{W}_h^Q, \boldsymbol{K}_h = \boldsymbol{H}^{(l-1)}\boldsymbol{W}_h^K$, and $\boldsymbol{V}_h = \boldsymbol{H}^{(l-1)}\boldsymbol{W}_h^V$ (Vaswani et al., 2017).

Finally, let $\boldsymbol{h}^{(L)}(\boldsymbol{s},t) \in \mathbb{R}^{d_{\mathrm{model}}}$ denote the output from the last transformer layer at spatio-temporal input $(\boldsymbol{s},t)$. The transformer-based approximation to the coefficient function is then obtained using a two-layer multilayer perceptron (MLP) regression head as

$$N(\boldsymbol{s},t \mid \boldsymbol{\theta}) = N_{\mathrm{out}}\Big(\boldsymbol{h}^{(L)}(\boldsymbol{s},t) \mid \boldsymbol{\theta}_{\mathrm{out}}\Big), \tag{4.2.5}$$

where $N_{\mathrm{out}}(\cdot \mid \boldsymbol{\theta}_{\mathrm{out}})$ denotes the output network. Under this formulation, the model learns dependence between space and time through the joint embedding together with self-attention, which makes it particularly suitable for coefficient functions with complex non-separable spatio-temporal patterns that may be difficult to characterize using simple architectures.

In the following, we use the general notation $N(\boldsymbol{s},t \mid \boldsymbol{\theta})$ to represent any of the spatio-temporal neural network formulations introduced above. To improve interpretability under the sparsity assumption on the effect functions, we impose a hard-thresholding operator at the output layer. For a vector $\boldsymbol{v}$ with threshold $\boldsymbol{\eta}$ and truncation level B, define

$$T_{\boldsymbol{\eta}}^{(B)}(\boldsymbol{v})_j = \mathrm{sign}(v_j) \cdot \mathbb{I}\{|v_j| \geq \eta_j\} \cdot \min(|v_j|, B),$$

where $B > 0$ is taken to be sufficiently large to ensure numerical stability (Zhang et al., 2023). In particular, the covariate-specific thresholds $\boldsymbol{\eta}_\beta \in \mathbb{R}_+^J$ and the subject-specific thresholds $\boldsymbol{\eta}_\alpha \in \mathbb{R}_+^N$ are introduced to impose sparsity directly at the level of the estimated coefficient functions, while preserving the spatio-temporally varying interpretation of $\boldsymbol{\beta}(\boldsymbol{s}, t)$ and $\boldsymbol{\alpha}(\boldsymbol{s}, t)$.

We then approximate the components in Model (4.2.1) by

$$\begin{aligned} \boldsymbol{\beta}(\boldsymbol{s}, t) &= T_{\boldsymbol{\eta}_\beta}^{(B)}\Big(N(\boldsymbol{s}, t \mid \boldsymbol{\theta}_\beta)\Big) \in \mathbb{R}^J, \\ \boldsymbol{\alpha}(\boldsymbol{s}, t) &= T_{\boldsymbol{\eta}_\alpha}^{(B)}\Big(N(\boldsymbol{s}, t \mid \boldsymbol{\theta}_\alpha)\Big) \in \mathbb{R}^N, \\ \log \sigma^2(\boldsymbol{s}, t) &= T_0^{(B)}\Big(N(\boldsymbol{s}, t \mid \boldsymbol{\theta}_\sigma)\Big) \in \mathbb{R}. \end{aligned} \tag{4.2.6}$$

Here,

$$\begin{aligned} N(\boldsymbol{s}, t \mid \boldsymbol{\theta}_\beta) &= \Big[N(\boldsymbol{s}, t \mid \boldsymbol{\theta}_{\beta_1}), \ldots, N(\boldsymbol{s}, t \mid \boldsymbol{\theta}_{\beta_J})\Big]^\top \\ N(\boldsymbol{s}, t \mid \boldsymbol{\theta}_\alpha) &= \Big[N(\boldsymbol{s}, t \mid \boldsymbol{\theta}_{\alpha_1}), \ldots, N(\boldsymbol{s}, t \mid \boldsymbol{\theta}_{\alpha_N})\Big]^\top, \end{aligned}$$

which are obtained by concatenating J neural networks for the main effects and N neural networks for the individual effects, respectively. Therefore, the hard-thresholding operator is applied componentwise at the output layer of each neural network representation.

The above model construction treats each spatio-temporal coordinate $(\boldsymbol{s}, t)$ as an input sample. As a result, the effective sample size becomes $V \times T$, which is typically large in imaging studies. This feature is particularly advantageous in neuroimaging, where the number of subjects is often limited while the number of observed spatial and temporal coordinates is substantial. Consequently, the DNN-based approximation strategy originally developed for purely spatial settings (Zhang et al., 2023) can be naturally adapted to the spatio-temporal domain.

4.2.3 Estimation Procedure

To estimate the main effect $\boldsymbol{\beta}(\boldsymbol{s}, t)$, the individual effect $\boldsymbol{\alpha}(\boldsymbol{s}, t)$, and the error variance function $\sigma^2(\boldsymbol{s}, t)$ in Model (4.2.1), we adopt a sequential three-step estimation procedure (Zhang et al., 2023).

In the first step, we estimate the main effect function $\boldsymbol{\beta}(\boldsymbol{s}, t) = \Big(\beta_1(\boldsymbol{s}, t), \ldots, \beta_J(\boldsymbol{s}, t)\Big)^\top$. Given the covariates $\boldsymbol{X}(t_k)$ and the image response $\boldsymbol{y}(\boldsymbol{s}_v, t_k)$ observed at time point t_k, $k = 1, \ldots, T$, and spatial location $\boldsymbol{s}_v$, $v = 1, \ldots, V$, we

estimate $\boldsymbol{\theta}_\beta$ by minimizing the spatio-temporally aggregated squared error

$$\widehat{\boldsymbol{\theta}}_\beta = \arg\min_{\boldsymbol{\theta}_\beta} \sum_{v=1}^{V} \sum_{k=1}^{T} \left\| \boldsymbol{y}(\boldsymbol{s}_v, t_k) - \boldsymbol{X}(t_k)\, T_{\boldsymbol{\eta}_\beta}^{(B)}\Big(N(\boldsymbol{s}_v, t_k \mid \boldsymbol{\theta}_\beta)\Big) \right\|_2^2. \tag{4.2.7}$$

The corresponding estimator of the main effect is then given by

$$\widehat{\boldsymbol{\beta}}(\boldsymbol{s}_v, t_k) = T_{\boldsymbol{\eta}_\beta}^{(B)}\Big(N(\boldsymbol{s}_v, t_k \mid \widehat{\boldsymbol{\theta}}_\beta)\Big).$$

With the estimated main effect $\widehat{\boldsymbol{\beta}}(\boldsymbol{s}_v, t_k)$ obtained from the first step, we next estimate the individual effect $\boldsymbol{\alpha}(\boldsymbol{s}, t)$ using the residualized responses. Specifically, we estimate $\boldsymbol{\theta}_\alpha$ by solving

$$\widehat{\boldsymbol{\theta}}_\alpha = \arg\min_{\boldsymbol{\theta}_\alpha} \sum_{v=1}^{V} \sum_{k=1}^{T} \left\| \boldsymbol{y}(\boldsymbol{s}_v, t_k) - \boldsymbol{X}(t_k)\widehat{\boldsymbol{\beta}}(\boldsymbol{s}_v, t_k) - T_{\boldsymbol{\eta}_\alpha}^{(B)}\Big(N(\boldsymbol{s}_v, t_k \mid \boldsymbol{\theta}_\alpha)\Big) \right\|_2^2. \tag{4.2.8}$$

Since $\boldsymbol{\alpha}(\boldsymbol{s}_v, t_k)$ is subject-specific, this optimization can be implemented separately for each individual. Equivalently, for subject i, we solve

$$\widehat{\boldsymbol{\theta}}_{\alpha_i} = \arg\min_{\boldsymbol{\theta}_{\alpha_i}} \sum_{v=1}^{V} \sum_{k=1}^{T} \left\| y_i(\boldsymbol{s}_v, t_k) - \boldsymbol{X}_i^\top(t_k)\widehat{\boldsymbol{\beta}}(\boldsymbol{s}_v, t_k) - T_{\eta_{\alpha_i}}^{(B)}\Big(N(\boldsymbol{s}_v, t_k \mid \boldsymbol{\theta}_{\alpha_i})\Big) \right\|_2^2. \tag{4.2.9}$$

Thus, this step involves fitting N neural networks, one for each individual effect function. The resulting estimator of the individual effect vector is

$$\widehat{\boldsymbol{\alpha}}(\boldsymbol{s}, t) = \Big[\widehat{\alpha}_1(\boldsymbol{s}, t), \ldots, \widehat{\alpha}_N(\boldsymbol{s}, t)\Big]^\top,$$

where

$$\widehat{\alpha}_i(\boldsymbol{s}_v, t_k) = T_{\eta_{\alpha_i}}^{(B)}\Big(N(\boldsymbol{s}_v, t_k \mid \widehat{\boldsymbol{\theta}}_{\alpha_i})\Big).$$

After obtaining both $\widehat{\boldsymbol{\beta}}(\boldsymbol{s}, t)$ and $\widehat{\boldsymbol{\alpha}}(\boldsymbol{s}, t)$, we proceed to estimate the error variance function $\sigma^2(\boldsymbol{s}, t)$. We begin by computing the mean squared residual

$$\bar{\sigma}^2(\boldsymbol{s}_v, t_k) = \frac{1}{N} \left\| \boldsymbol{y}(\boldsymbol{s}_v, t_k) - \boldsymbol{X}(t_k)\widehat{\boldsymbol{\beta}}(\boldsymbol{s}_v, t_k) - \widehat{\boldsymbol{\alpha}}(\boldsymbol{s}_v, t_k) \right\|_2^2.$$

We then estimate the log-scale variance function by solving

$$\widehat{\boldsymbol{\theta}}_\sigma = \arg\min_{\boldsymbol{\theta}_\sigma} \sum_{v=1}^{V} \sum_{k=1}^{T} \left\| \log \bar{\sigma}^2(\boldsymbol{s}_v, t_k) - N(\boldsymbol{s}_v, t_k \mid \boldsymbol{\theta}_\sigma) \right\|_2^2. \tag{4.2.10}$$

The resulting estimator of the error variance is given by

$$\widehat{\sigma}^2(\boldsymbol{s}_v, t_k) = \exp\left\{T_0^{(B)}\Big(N(\boldsymbol{s}_v, t_k \mid \widehat{\boldsymbol{\theta}}_\sigma)\Big)\right\}.$$

Finally, the thresholds $\boldsymbol{\eta}_\beta$ and $\boldsymbol{\eta}_\alpha$ are selected using a random decoupling procedure that approximates the null distribution of the coefficient estimates (Barut et al., 2016; Zhang et al., 2023). Specifically, the subject labels in $\boldsymbol{X}$ are repeatedly permuted to generate null estimates using the three-step estimation procedure described above, and the threshold for each component is chosen as the 0.95 quantiles of the corresponding absolute null estimates.

4.3 Numerical Studies

In this section, we evaluate the estimation performance of the proposed methods for spatio-temporally varying coefficients under four representative simulation settings: (i) a non-sparse additive spatio-temporal pattern, (ii) a sparse multiplicative spatio-temporal pattern, (iii) a sparse concatenative spatio-temporal pattern, and (iv) a sparse concatenative spatio-temporal pattern with a moving activation region. These settings are designed to reflect different levels of structural complexity and to assess how well the proposed methods adapt to both separable and non-separable spatio-temporal dependence.

Throughout the simulations, the spatial coordinate is taken as $\boldsymbol{s}_v \in \mathcal{S} \subset [0,1]^D$ on a regular grid of V locations with 64×64 regions. The temporal domain is set to $[0,1]$ and is discretized into $T = 100$ equally spaced time points. For simplicity, we consider a univariate covariate generated from $N(0,1)$. For $N = 20$ subjects, the corresponding image response is generated according to model (4.2.1). The spatial patterns are adapted from the image response regression simulations in Zhang et al. (2023) and are then combined with smooth or piecewise-smooth temporal components to construct the spatio-temporal design studied here. In particular, the main effect templates contain activation regions with both sharp and smooth boundaries, the individual effect templates follow cross-shaped activation patterns, and the log-variance function is generated from a truncated sinusoidal pattern. Building on these spatial templates, we introduce temporal variation in different ways to construct the four simulation scenarios described below.

Following the general strategy in Zhang et al. (2023), we evaluate the estimation accuracy of the main effect at each time point using the relative mean squared error,

computed as

$$\text{Relative MSE}(t) = \frac{\sum_{j=1}^{J}\sum_{v=1}^{V}\left[\widehat{\beta}_j(\boldsymbol{s}_v,t) - \beta_j(\boldsymbol{s}_v,t)\right]^2/(VJ)}{\sum_{j=1}^{J}\sum_{v=1}^{V}\beta_j^2(\boldsymbol{s}_v,t)/(VJ)}, \tag{4.3.1}$$

which measures the relative estimation error at time t. We report this quantity over time to assess how estimation accuracy changes across the temporal domain.

For comparison, we also apply IRRNN method of Zhang et al. (2023) separately at each time point. Since that method is designed for spatial coefficient functions and does not explicitly model temporal dependence, we further consider a local temporal aggregation strategy. Specifically, for a target time point t_k, we pool the images from the w preceding and w following time points and use the aggregated information to estimate the coefficient function at t_k. In the experiments, we consider three window sizes, $w = 3$, 5, and 10, to assess how the amount of local temporal information affects the performance of the baseline IRRNN method.

4.3.1 Spatial Design

For the main effect function $\beta_j(\boldsymbol{s}_v, t)$, we consider spatial activation template with both sharp and smooth boundaries, following the general simulation framework of Zhang et al. (2023). These patterns serve as the spatial templates and are subsequently combined with different temporal functions to generate four spatio-temporal scenarios.

For activation regions with sharp boundaries, we generate square-shaped regions with piecewise smooth values inside each square. The top-left corner of a square is determined by a random location (u_1, u_2), where u_1 and u_2 are independently sampled from Uniform$(0, 0.6)$. The side length of the square is set to 0.4, ensuring that the entire region lies within the unit domain. Let (c_1, c_2) denote the center of the square. For a spatial location $\boldsymbol{s}_v$ inside the square, the activation magnitude is defined by the Euclidean distance to the center as $\beta_j(\boldsymbol{s}_v) = \sqrt{(s_{v1} - c_1)^2 + (s_{v2} - c_2)^2}$. To introduce variability in the spatial profile, we randomly choose whether the activation values increase or decrease with distance from the center using a Bernoulli(0.5) variable. This pattern is truncated at its 60th percentile, so that any value above this threshold is set equal to the threshold. This operation produces a spatial surface with a flat outer plateau and a sharp boundary. Finally, the truncated values are rescaled to the interval $[0, 1]$, and then mapped to the range $[0.5a, a]$, where $a \sim$ Uniform$(1, 2)$, with a random sign in $\{-1, 1\}$ assigned to the entire activation region.

For activation regions with smooth boundaries, we generate circular patterns with radius $r = 0.5$ and center (c_1, c_2), where c_1 and c_2 are independently sam-

pled from Uniform$(0.25, 0.75)$. For a spatial location $\boldsymbol{s}_v = (s_{v1}, s_{v2})$ inside the circle, the base value is determined by its Euclidean distance to the center as $b(\boldsymbol{s}_v) = \sqrt{[(s_{v1} - c_1)/r\ ]^2 + \ [(s_{v2} - c_2)/r\ ]^2}$. The resulting activation values are min-max normalized and then linearly mapped as $(1-b)\times a$ where $a \sim$ Uniform$(7, 8)$, with a random sign in $\{-1, 1\}$ assigned to the entire activation region. This construction yields a smoothly varying spatial pattern that decreases from the center toward the boundary of the circle.

The sharp- and smooth-boundary activation regions are then combined on the spatial domain to form the main effect template. This combined spatial pattern serves as the basis for each main effect function. We denote the resulting spatial template for the main effect by $f_\beta(\boldsymbol{s}_v)$. An illustration of the spatial design is given in Figure C.1a.

We next describe the spatial pattern used for the individual effect, which is generated as a cross-shaped activation region with an underlying radial profile. We first generate a radial pattern by sampling a center (c_1, c_2), where $c_1, c_2 \sim$ Uniform$(0, 1)$, together with a mean level from Uniform$(0, 1)$. For a spatial location $\boldsymbol{s}_v = (s_{v1}, s_{v2})$, the base value is defined by its Euclidean distance to the center as $\sqrt{(s_{v1} - c_1)^2 + (s_{v2} - c_2)^2} \times 2$ and is then centered so that the resulting radial pattern has mean zero. Based on this radial pattern, we construct a cross-shaped activation region by selecting the left edge of the vertical bar and the top edge of the horizontal bar independently from Uniform$(0, 0.8)$, with both bar widths fixed at 0.2. Values outside the cross-shaped region are set to zero, and values inside the region inherit the centered radial pattern. To preserve the zero-centered structure of the active region, the final activation values are recentered using only the nonzero values within the cross. We repeat this procedure independently for each subject, so that the individual effect functions vary across subjects in both locations and intensities of the cross-shaped activation region. We denote the resulting individual-effect spatial template for subject i by $f_{\alpha_i}(\boldsymbol{s}_v)$. An illustration of the individual-effect spatial design is given in Figure C.1b.

For the log variance function of the error term, we generate a sinusoidal spatial pattern. Specifically, for a spatial location $\boldsymbol{s}_v = (s_{v1}, s_{v2})$, we set $\log(\sigma^2(\boldsymbol{s}_v)) = \sum_{d=1}^{2} \sin\Big(4\pi(s_{vd} - a_d)\Big)$, where $a_1, a_2 \sim$ Uniform$(0, 1)$. This construction yields a spatially varying variance with periodic oscillations over the two-dimensional domain. To induce sparsity in the variance pattern, we truncate the values in an intermediate range to zero. In particular, all values lying between $0.7\min\{\log\sigma^2(\boldsymbol{s}_v)\}$ and $0.3\max\{\log\sigma^2(\boldsymbol{s}_v)\}$ are truncated to 0. We denote the resulting spatial template for the log-variance function by $f_\sigma(\boldsymbol{s}_v)$. An illustration of the variance design is given in Figure C.1c.

4.3.2 Temporal Design

Given the spatial templates introduced above, we incorporate temporal variation through a time-dependent function, denoted by $f_t(t)$. This temporal component determines how the magnitude and sign of the signal evolve over time, and allows the resulting spatio-temporal patterns to exhibit gradual amplification, attenuation, oscillation, or sign changes across the temporal domain.

In the following spatio-temporal simulation scenarios, we consider several representative temporal mechanisms, including linear, power, sinusoidal, and piecewise-varying trends, and combine them with the main-effect spatial templates in different ways. These constructions yield a range of spatio-temporal patterns with varying levels of complexity, including both separable and non-separable structures within the main effect, thereby facilitating comparison across different scenarios. For ease of presentation, we introduce the temporal component here in a generic form and defer the specific choices for each setting to the corresponding subsections.

4.3.3 Spatio-Temporal Design

Case I: Additive Spatio-temporal Design We start with a simple scenario in which the main effect follows a separable additive spatio-temporal structure with a sharp-boundary spatial pattern. Denote the spatial template as $f_{\boldsymbol{\beta}}(\boldsymbol{s}_v)$ and define the temporal function as $f_t(t_k) = t_k - 0.5$ for $t_k \in [0, 1]$. The resulting main effect function is given by $\beta_1(\boldsymbol{s}_v, t_k) = f_{\boldsymbol{\beta}}(\boldsymbol{s}_v) + f_t(t_k)$. This construction yields an additive spatio-temporal pattern in which the spatial activation map remains fixed, while the overall signal level varies linearly over time. An example of this design is shown in Figure C.2a.

Case II: Multiplicative Spatio-temporal Design We next consider a separable multiplicative spatio-temporal setting in which the main effect is constructed from the sharp-boundary spatial template $f_\beta(\boldsymbol{s}_v)$ and a sinusoidal temporal component. Here, we choose the temporal trend to be $f_t(t_k) = \sin(2\pi t_k + 1.5)/2.5$ for $t_k \in [0, 1]$. With the spatial design $f_{\boldsymbol{\beta}}(\boldsymbol{s}_v)$, the main effect function is computed as $\beta_1(\boldsymbol{s}_v, t_k) = |f_{\boldsymbol{\beta}}(\boldsymbol{s}_v)| \times f_t(t_k)$. This construction yields a sparse multiplicative spatio-temporal pattern in which the absolute spatial activation map is modulated over time by a sinusoidal trend for nonzero values. An example of this design is shown in Figure C.2b.

Case III: Concatenative Spatio-temporal Design We then consider a non-separable spatio-temporal design that contains both sharp- and smooth-boundary activation patterns. In this design, we consider a linearly increasing temporal trend

$f_t^1(t_k) = t_k$ and a linearly decreasing temporal trend $f_t^2(t_k) = 1 - t_k$ for $t_k \in [0, 1]$. The spatial template consists of two patterns: a smooth circular pattern $f_{\boldsymbol{\beta}}^1(\boldsymbol{s}_{v_1})$ on a subset of locations $\boldsymbol{s}_{v_1}$ and a sharp square pattern $f_{\boldsymbol{\beta}}^2(\boldsymbol{s}_{v_2})$ on another subset of locations $\boldsymbol{s}_{v_2}$. To generate non-separable spatio-temporal effects, we employ an increasing temporal trend on the absolute activation within the smooth circular regions by $|f_{\boldsymbol{\beta}}^1(\boldsymbol{s}_{v_1})| \times f_t^1(t_k)$ and a decreasing temporal trend on the absolute activation within the sharp square regions by $|f_{\boldsymbol{\beta}}^2(\boldsymbol{s}_{v_2})| \times f_t^2(t_k)$. These two patterns are combined in the same image shape as in the previous designs. The remaining locations remain inactive throughout the time period, thereby preserving sparsity in the main-effect design. An example of this design is shown in Figure C.2c.

Case IV: Concatenative Spatio-temporal Design with a Moving Activation Region Finally, we consider a more challenging non-separable spatio-temporal design with a moving activation region. As in Case III, the main effect combines a sharp-boundary square component and a smooth circular component, but here the circular pattern varies in both intensity and spatial location over time. Specifically, we apply a linearly increasing temporal trend $f_t^1(t_k) = 0.2 + 1.3t_k$ to a sharp square $f_{\boldsymbol{\beta}}^1(\boldsymbol{s}_{v_1})$ by $|f_{\boldsymbol{\beta}}^1(\boldsymbol{s}_{v_1})| \times f_t^1(t_k)$ and a linearly decreasing temporal trend $f_t^2(t_k) = 1 - t_k$ on a smooth circular pattern by $|f_{\boldsymbol{\beta}}^2(\boldsymbol{s}_{v_2})| \times f_t^2(t_k)$ for $t_k \in [0, 1]$. In addition to this concatenative spatio-temporal pattern, we further update the spatial location of the smooth circular pattern across time. In particular, the center of the circle c_{t_k} at time t_k is set to $c_{t_k} = (1 - t_k)c_{\text{start}} + t_k c_{\text{end}}$, where c_{start} and c_{end} denote the start and end centers of the moving circle. The absolute activation values within the circle are updated as $|f_{\boldsymbol{\beta}}^2(\boldsymbol{s}_{v_2})| \times f_t^2(t_k) \times \max(0, 1 - d_{t_k}(\boldsymbol{s}_v)/r)$ for the Euclidean distance from a spatial location $\boldsymbol{s}_v$ to the center $d_{t_k}(\boldsymbol{s}_v) = \|\boldsymbol{s}_v - c_{t_k}\|_2$. These two patterns are then combined together in the same image domain as in the previous designs, leaving the remaining locations inactive to preserve sparsity in the main-effect design. Consequently, both the amplitude and the spatial location of the main effect vary over time, making this setting more challenging than the previous cases. An example of this design is shown in Figure C.2d.

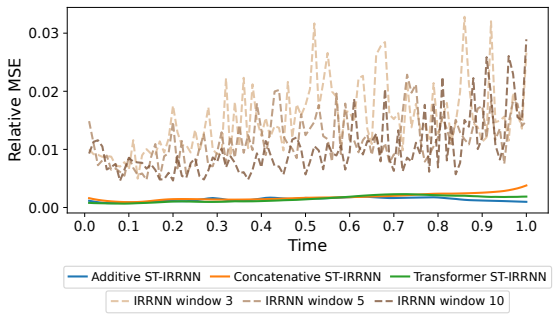
Across all four simulation cases, the spatio-temporal designs described above are imposed on the main effect. To isolate the role of the main-effect structure, we do not introduce additional temporal variation in the individual effect or the error variance. For each scenario, we evaluate the proposed ST-IRRNN methods and compare them with the IRRNN method from Zhang et al. (2023). For the baseline method, IRRNN is fitted separately at each time point t_k, and local temporal information is incorporated by aggregating images within a temporal window of size $w \in \{3, 5, 10\}$ around t_k. In all four settings, we apply the concatenative ST-IRRNN (4.2.2) and

the transformer-based ST-IRRNN (4.2.5), since both are designed to capture general spatio-temporal dependence. In addition, under the separable spatio-temporal scenarios, we apply the additive ST-IRRNN (4.2.3) in the additive spatio-temporal setting and the multiplicative ST-IRRNN (4.2.4) in the multiplicative spatio-temporal setting. Because the additive design does not impose sparsity on the main effect, we do not apply thresholding to any of the methods in the additive setting.

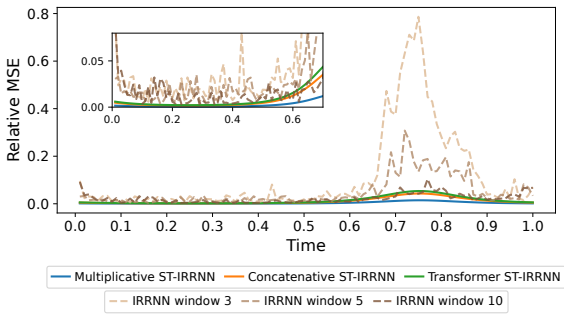
Figure 4.1 displays the relative MSE (4.3.1) averaged over 20 replications for all methods under the four spatio-temporal simulation settings. Overall, the proposed ST-IRRNN methods consistently achieve smaller estimation error and exhibit smoother temporal trends than the IRRNN baselines with fixed time windows. This pattern is observed across all four scenarios, suggesting that explicitly modeling the spatio-temporal dependence tends to provide better estimation performance than applying a spatial method separately at each time point with local temporal aggregation. In the first two cases, where the true spatio-temporal structure is separable, the ST-IRRNN variant matched to the underlying design performs best. In Case I, the additive ST-IRRNN attains the lowest and most stable relative MSE over time, and in Case II the multiplicative ST-IRRNN returns the most accurate estimation of the true main effect. This behavior is consistent with the underlying spatio-temporal design and indicates that incorporating the correct structural form can improve estimation accuracy. In Cases III and IV, where the main effect is non-separable and the temporal dynamics are more complex, both the concatenative and transformer-based ST-IRRNN methods continue to outperform the fixed-window IRRNN baselines. In these two settings, the transformer-based ST-IRRNN generally yields slightly smaller error and smoother temporal behavior than the concatenative ST-IRRNN. This suggests that the transformer architecture is better able to accommodate more complex and dynamically varying spatio-temporal dependence. Meanwhile, the strong performance of the concatenative ST-IRRNN indicates that directly learning from the joint spatio-temporal input is already highly effective when the underlying effect is non-separable. Overall, these results suggest that explicitly learning complex spatio-temporal dependence through concatenation or a transformer-based ST-IRRNN approach can help improve both estimation accuracy and stability.

For the fixed-window IRRNN baselines, larger window sizes generally lead to lower error and reduced variability throughout time, indicating that incorporating additional local temporal information can partially improve its performance. Nevertheless, all the IRRNN baselines show consistently higher relative MSE and greater variability throughout time. Taken together, these results show that explicit spatio-temporal modeling is important for accurately estimating dynamically varying coefficient functions, particularly when the underlying effects are non-separable or evolve through

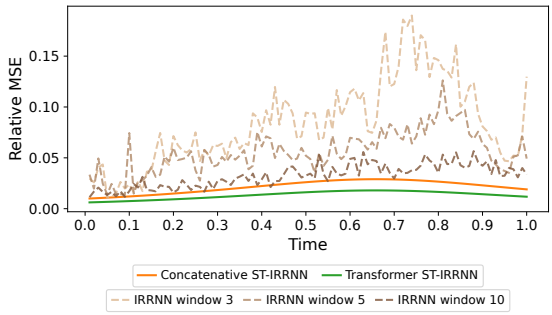
more complex mechanisms over time.



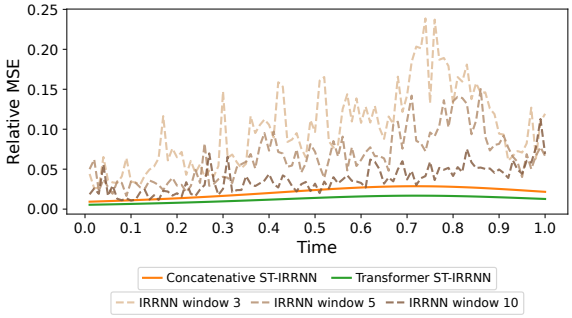
(a) Case I: Additive spatio-temporal design



(b) Case II: Multiplicative spatio-temporal design



(c) Case III: Concatenative spatio-temporal design



(d) Case IV: Concatenative spatio-temporal design with a moving activation region

Figure 4.1: Relative MSE over time for the four spatio-temporal simulation settings: (a) additive, (b) multiplicative, (c) concatenative, and (d) concatenative with a moving activation region. In panel (a), the solid curves correspond to additive, concatenative, and transformer-based ST-IRRNN. In panel (b), they correspond to multiplicative, concatenative, and transformer-based ST-IRRNN. In panels (c) and (d), they correspond to concatenative and transformer-based ST-IRRNN. The dashed curves represent the fixed-window IRRNN with temporal window sizes $\{3, 5, 10\}$.

4.4 Data Application

Continuous fMRI acquisitions under naturalistic paradigms have become increasingly common in neuroimaging, as they allow researchers to track brain responses over time under dynamic stimuli such as movie viewing. Such audio-visual stimuli are multimodal, context-rich, and evolving over time. They provide an ecologically valid framework for studying brain function in settings that more closely approximate real-world experience (Sonkusare et al., 2019; Finn and Bandettini, 2021; Matusz et al., 2019). Previous studies have shown that naturalistic stimuli involve distributed brain systems, induce synchronized responses across viewers (Hasson et al., 2004), reveal

hierarchical processing across multiple temporal receptive windows (Hasson et al., 2008; Honey et al., 2012; Lerner et al., 2011), and modulate large-scale network organization during narrative comprehension (Simony et al., 2016). These processes are closely related to higher-order functions such as social cognition, event perception, and narrative understanding (Lahnakoski et al., 2012; Baldassano et al., 2017). Therefore, examining effects from scene-dependent stimuli in such data can provide useful insight into how different forms of naturalistic information dynamically engage large-scale brain networks.

In this section, we apply the proposed ST-IRRNN approaches to an fMRI dataset collected during continuous audio-visual stimulation in order to investigate scene-dependent spatio-temporal patterns of brain responses (Chen et al., 2017). This dataset consists of fMRI measurements from 17 subjects who watched the first 50 minutes of the BBC television series Sherlock. The movie was presented in two scans and segmented into a total of 50 movie scenes, each associated with 10 semantic variables. These fMRI measurements were taken every 1.5 seconds of the movie, yielding $1,976$ scans each with dimension $61 \times 73 \times 61$. The imaging data were pre-processed using slice-timing correction, motion correction, linear detrending, high-pass filtering, coregistration and affine transformation (Chen et al., 2017). In the present analysis, we focus on the images collected during the first movie scan, which contains 945 images spanning 27 scenes. To make the analysis computationally manageable, we apply the Schaefer2018 parcellation (Schaefer et al., 2018) and obtain $1,000$ regions of interest (ROIs), each represented by a three-dimensional coordinate rescaled to $[-1, 1]$. In this study, we are primarily interested in the effects of audio, visual, and emotion-related stimuli, which serve as three complementary aspects of scene content. Accordingly, we consider three scene-level covariates from Chen et al. (2017): (i) the proportion of time during a scene in which anyone is speaking, (ii) the number of people on screen, and (iii) an arousal level reflecting the level of excitement, engagement, or activity assigned by independent raters.

Because the underlying spatio-temporal structure is unknown in this real-data application, we focus on the concatenative and transformer-based ST-IRRNN methods, both of which are designed to capture general spatio-temporal dependence without requiring a prespecified separable form. As a benchmark, we compare their performance with the baseline IRRNN method using a fixed temporal window of size 100. For model fitting and evaluation, images from 12 randomly selected subjects are used for training, and the remaining 5 subjects are reserved for testing. Hyperparameters are selected by three-fold cross-validation. Because the subjects in the training and test sets are different, the individual effects estimated from the training data cannot be directly transferred to unseen subjects. Therefore, we focus on the main effects

associated with the three scene-level covariates, which are intended to capture audio, visual, and emotion-related stimuli and their associations with neural activity.

Under this setting, prediction on the test set is based on the estimated main effects only. For a subject i, the predicted image response at spatial location $\boldsymbol{s}_v$ and time t is given by $\widehat{y}_i(\boldsymbol{s}_v, t) = \sum_{j=1}^{J} X_{ij}(t)\widehat{\beta}_j(\boldsymbol{s}_v, t)$. Based on this prediction, we evaluate performance at each time point using the relative test MSE

$$\frac{\sum_{i=1}^{N}\sum_{v=1}^{V}[\widehat{y}_i(\boldsymbol{s}_v, t) - y_i(\boldsymbol{s}_v, t)\ ]^2/(NV)}{\sum_{i=1}^{N}\sum_{v=1}^{V} y_i^2(\boldsymbol{s}_v, t)/(NV)},$$

which measures the discrepancy between the main-effect-based prediction and the observed image response at each time point, normalized by the overall signal magnitude. The results are reported in Figure 4.2.

Figure 4.2 shows that both ST-IRRNN approaches generally achieve lower relative test MSE than the IRRNN baseline for most time frames, indicating that incorporating temporal input in spatio-temporal dependence improves prediction of the observed image responses from the scene-dependent covariates. Although the improvement is not uniform at every time point, the advantage of the ST-IRRNN methods is persistent over a large proportion of the movie scan. In contrast, the IRRNN baseline produces a relatively flatter trajectory, but remains at a consistently higher error level throughout most of the time domain. Meanwhile, the two ST-IRRNN curves display more fluctuations, with occasional spikes above the baseline. Such behavior is not surprising, because predictions are constructed only from the estimated main effects, while the observed fMRI responses still contain subject-specific and temporal variation that is not directly modeled in the prediction step. Between the two proposed methods, the concatenative and transformer-based ST-IRRNN exhibit broadly similar temporal patterns, but the transformer-based approach appears slightly more stable and yields lower error over many intervals. Since the true spatio-temporal structure is unknown in this real-data analysis, these results should be interpreted primarily as predictive evidence. Taken together, these results suggest that incorporating spatio-temporal structure through ST-IRRNN provides a practical advantage for explaining image responses under continuous stimuli, even when the underlying structure is unknown.

We next examine the estimated main-effect coefficient maps to explore how the three covariates are associated with scene-dependent response patterns within brain systems. For visualization, Figure 4.3 displays the estimated coefficient maps at the middle time frame of six selected movie scenes. Across all six scenes, the concatenative and transformer-based ST-IRRNNs recover broadly consistent spatial patterns, which is in line with their similar predictive performance in Figure 4.2.

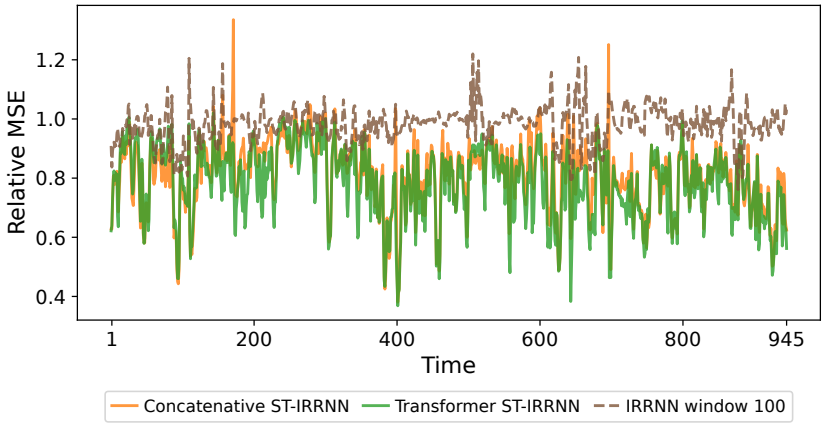


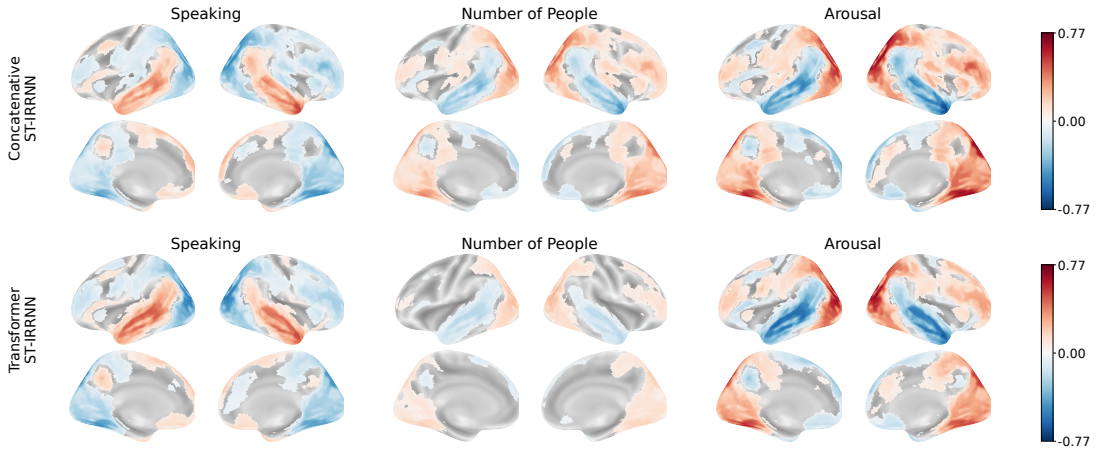
Figure 4.2: Relative test MSE over time for concatenative ST-IRRNN, transformer-based ST-IRRNN, and IRRNN with a time window size of 100 when predicting observed image responses from the estimated main effects in the Sherlock movie-watching data.

The estimated coefficient maps from both methods also show clear scene-specific and covariate-specific patterns, suggesting that the models are able to capture dynamic stimuli within each scene's content. In the Press conference scene (Figure 4.3b), the number of people effect is dominant, while the speaking and arousal effects are comparatively weak. In contrast, the Poison vial # 3 scene (Figure 4.3c) is mainly driven by the speaking effect, with smaller contributions from the other two covariates. The Sally didn't make it home scene (Figure 4.3f) is dominated by the arousal effect, which exhibits the strongest and most spatially extensive pattern in this panel. By comparison, Watson psychiatrist (Figure 4.3a), Watson and SH meet (Figure 4.3d), and Science of deduction conversation (Figure 4.3e) show mixed profiles in which more than one covariate contributes substantially to the image response without a clear dominant covariate effect. Overall, the estimated coefficients vary meaningfully with scene content, and the covariate pattern within each scene remains largely stable across the two ST-IRRNN architectures.

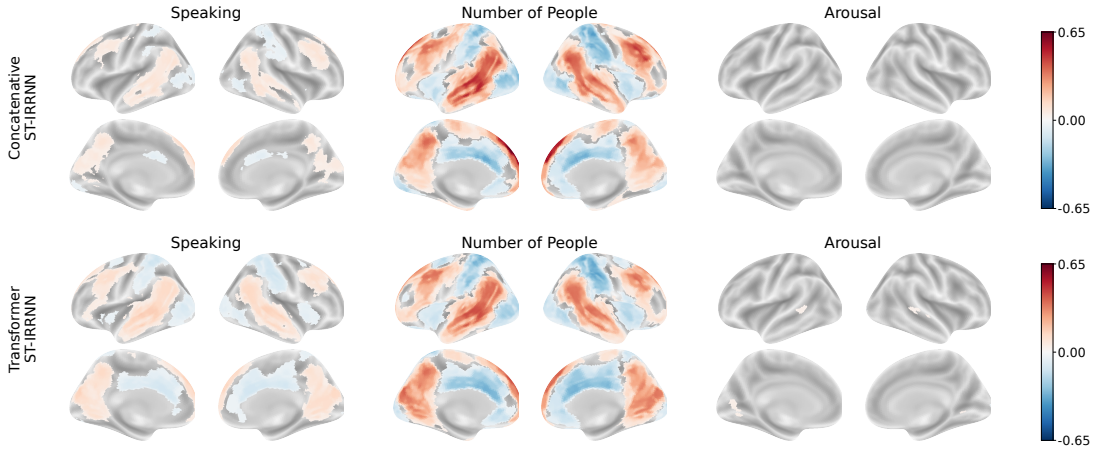
These scene-dependent coefficient patterns also provide a useful basis for interpreting how different naturalistic stimuli engage large-scale brain systems. At the brain network level (Yeo et al., 2011), the spatial distribution of these estimated effects is also broadly consistent with neurobiological interpretations. In the Press conference scene, which can be characterized by a complex visual environment with a high count of background characters, the number of people effect mainly involves regions in visual, dorsal attention, and ventral attention networks, with weaker extension into the default and somatomotor systems. This pattern within this social scene is broadly consistent with previous work showing that naturalistic social scenes recruit distributed systems for social perception and that dorsal and ventral attention net-

works are involved in orienting toward behaviorally relevant social cues (Lahnakoski et al., 2012; McMahon and Isik, 2023; Callejas et al., 2014; Corbetta and Shulman, 2002). During social interactions, the number of people is an important factor of neural responses (Martynenko et al., 2025). For sensory-dominant scenes, such as Poison vial # 3, the speaking covariate displays strong effects, while Watson psychiatrist and Watson and SH meet show more mixed profiles in which speaking and arousal both contribute substantially. In these scenes, the speaking effects are most evident in somatomotor, ventral attention, frontoparietal, and default networks, which are broadly compatible with auditory and language-related processing regions (Yeo et al., 2011). These active patterns align with evidence that these networks are critical for integrating auditory stimuli and semantic representations during speech perception and narrative comprehension (Liebenthal and Möttönen, 2018; Pulvermüller and Fadiga, 2010; Huth et al., 2016; Simony et al., 2016).

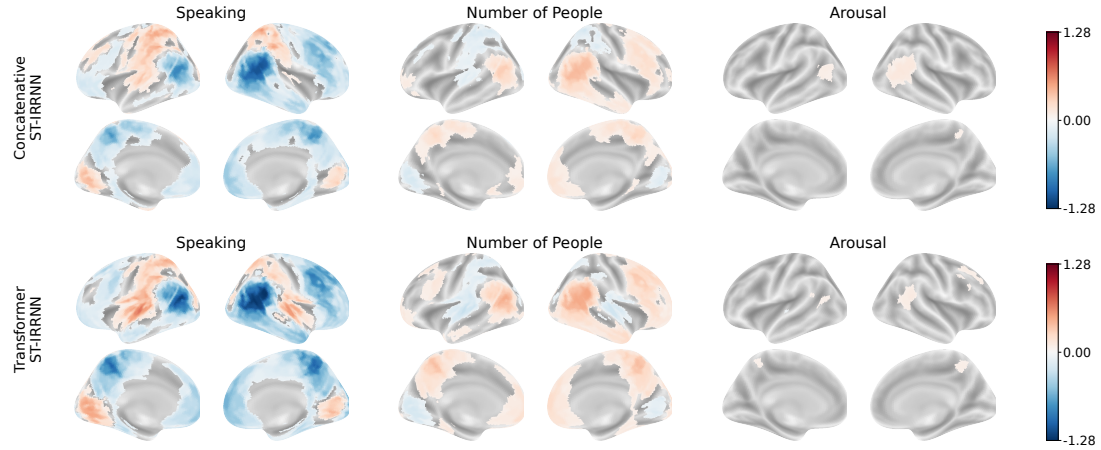
Beyond these sensory- and speech-dominant scenes, the Science of deduction and Watson and SH meet scenes show broad engagement of frontoparietal network, together with concurrent number of people and arousal effects, effectively capturing the dynamic involvement of cognitive networks during complex narrative transitions and cognitive demand. This pattern aligns with prior work that the frontoparietal system is closely related to cognitive control, integrating external cues with logical reasoning to track complex deductions (Vincent et al., 2008; Zanto and Gazzaley, 2013; Dosenbach et al., 2008; Bertolero et al., 2018). Finally, the arousal maps, especially in Sally didn't make it home and also in Science of deduction conversation scenes, are spatially extensive and span ventral attention, default, frontoparietal, and limbic networks. The strong effects from these networks match with studies showing that emotional and affective intensity are associated with multiple large-scale systems (Nummenmaa et al., 2012; Pessoa, 2008), including ventral attention, default, and frontoparietal networks (Satpute and Lindquist, 2019; Viviani, 2013; Ke et al., 2025), and can even be tracked by ventral attention and default networks (Lettieri et al., 2022). Overall, although these coefficient maps should be interpreted as descriptive model-based profiles rather than formal voxelwise localization results, they provide interpretations of how scene-level covariate effects vary over time and across large-scale brain systems.



(a) Watson psychiatrist scene

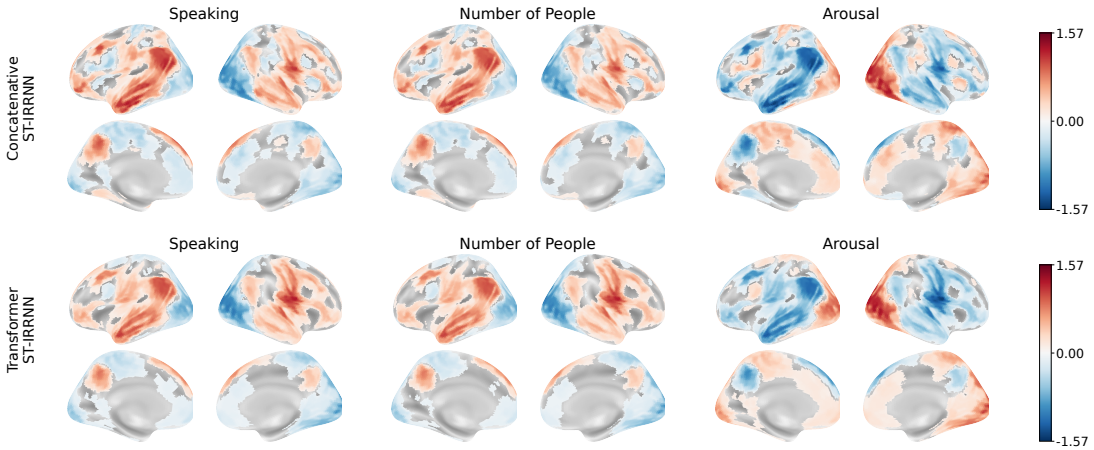


(b) Press conference scene

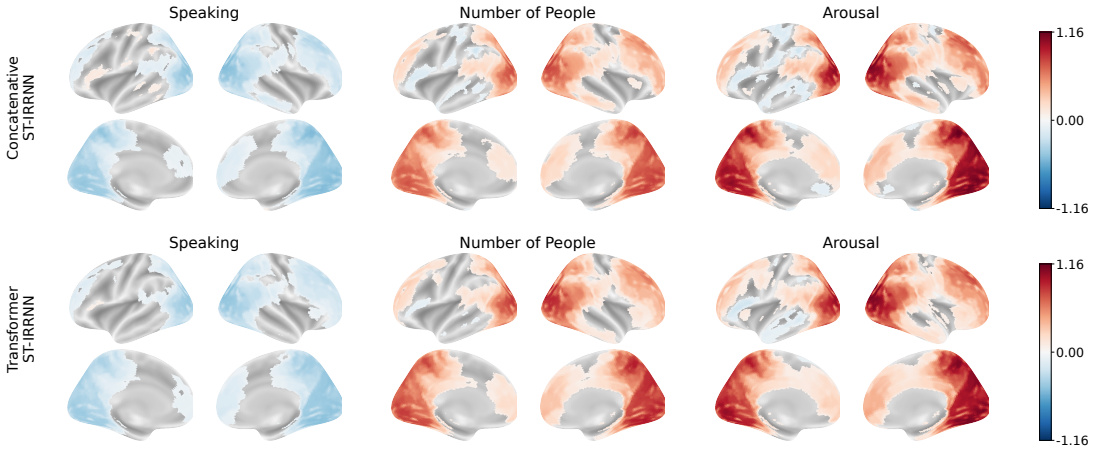


(c) Poison vial #3 scene

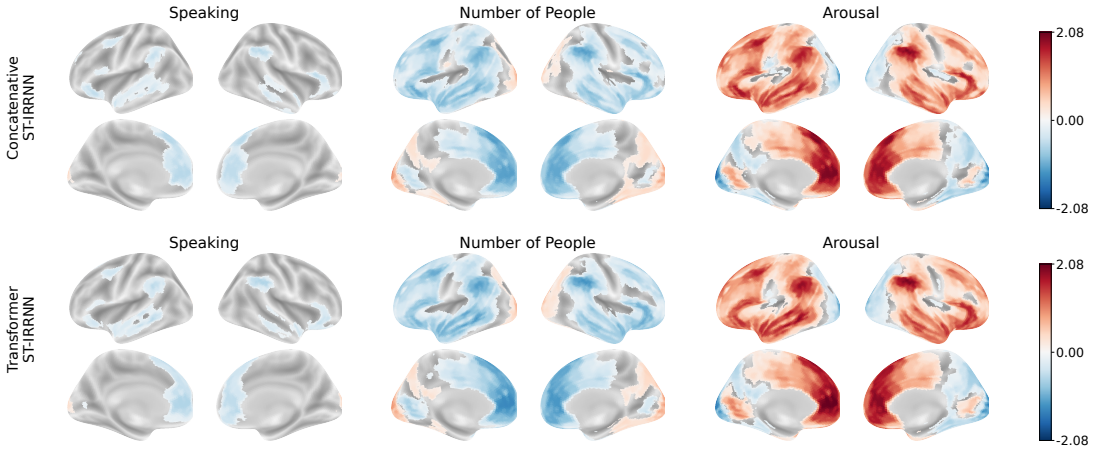
Figure 4.3: Covariate effect estimates from the concatenative ST-IRRNN (top row) and transformer-based ST-IRRNN (bottom row) for selected movie scenes. The three columns correspond to the speaking, number of people, and arousal covariates. Panels (a)–(c) display the Watson psychiatrist, Press conference, and Poison vial #3 scenes, respectively.



(d) Watson and SH meet scene



(e) Science of deduction conversation scene



(f) Sally did not make it home scene

Figure 4.3: Covariate effect estimates from the concatenative ST-IRRNN (top row) and transformer-based ST-IRRNN (bottom row) for selected movie scenes, continued. The three columns correspond to the speaking, number of people, and arousal covariates. Panels (d)–(f) display the Watson and SH meet, Science of deduction conversation, and Sally did not make it home scenes, respectively.

CHAPTER 5

Conclusion

This dissertation develops statistical methodologies designed to address the inherent complexities of neuroimaging data. The central goal of this dissertation is to construct models and procedures that are flexible to accommodate the high dimensionality, heterogeneity, and complex structures of neuroimaging data, while remaining interpretable from a scientific perspective. From the functional connectivity perspective, we introduce a ceFC metric as an unbiased measure of covariation between pairs of edges and develop consistent estimators in both general and high-dimensional settings. We also propose a multiple hypothesis testing framework for examining nonzero ceFC measures among edge pairs. Building upon this metric, we evaluate the practical utility of ceFC within a brain-wide association study procedure, demonstrating its limited incremental predictive value for phenotypic traits when a sufficiently rich set of conventional node-centric features is included. Finally, we shift from static connectivity structures to dynamic processes. We introduce ST-IRRNN methods to model spatially and temporally varying brain responses under continuous naturalistic stimuli, thereby effectively capturing and interpreting the effects of dynamic stimuli across large-scale brain systems.

Despite the advancements discussed here, there are several directions in which this work can be further developed. For the ceFC framework, while our thresholded estimator provides statistical advantages in high-dimensional settings, it shares with other sparsity-based methods the potential removal of biologically meaningful weak connections. Previous studies have highlighted the functional importance of weak ties in brain networks (Park and Friston, 2013; Gallos et al., 2012; Varley et al., 2023). However, with limited sample sizes typical in fMRI studies, it becomes statistically challenging to reliably distinguish genuinely weak biological signals from random noise across the $O(p^4)$ potential edge-edge connections. Our thresholding framework prioritizes the identification of robust, reproducible higher-order interactions while controlling false discoveries, though we recognize that some subtle yet important connections might be excluded. Future studies with larger sample sizes

or multi-site datasets may enable more powerful detection of subtle yet meaningful higher-order interactions.

Moreover, although ceFC effectively captures the higher-order connectivity in brain networks, it is limited in revealing the linear connections between edges. Considering the intricate and non-trivial structures inherent in brain networks and nervous systems (Stephan et al., 2008; Spiegler et al., 2011; Roberts and Robinson, 2012; Lahaye et al., 2003), such a linear measurement may not fully capture the complex non-linear patterns that could enhance sensitivity analysis across groups of people (Motlaghian et al., 2022; Kinsey et al., 2024). Therefore, it is meaningful to explore additional metrics that measure non-linear relationships among edges, such as distance correlation (Székely et al., 2007) for dependency measurement, providing a more comprehensive understanding of brain organizations.

The application of higher-order relationships in the analysis of brain networks provides new insights into functional brain organizations. Our study so far focuses on analyzing connectivity patterns within each subject without extending the analysis to include the comparison across a group of subjects, a limitation that does not account for potentially shared patterns or inter-subject variability. Looking forward, one promising direction includes the multi-subject analysis, exploring both the group-wise pattern among multiple subjects and the individual-specific characteristics of each subject. Understanding group-wise patterns could assist in disease diagnosis *via* the identification of biomarkers, and examining the individual-specific characteristics could provide insights into personalized medicine through personal information, extending current edge-based studies of group-wise comparison across healthy people and patients (Dai et al., 2023; Yang et al., 2023; Zamani Esfahlani et al., 2022; Sun et al., 2023) and individual differences (Ville et al., 2021).

For the spatio-temporal varying coefficient framework in Chapter 4, several methodological extensions are possible, including uncertainty quantification for the estimated coefficient functions and statistical inference for dynamic activation maps. It would also be valuable to investigate broader classes of naturalistic stimuli and additional forms of temporal representation, especially in settings where multiple stimulus features interact.

In conclusion, this dissertation addresses three related questions: how to characterize higher-order dependence in functional connectivity, whether such higher-order structure provides additional predictive value for phenotypic traits, and how to model dynamic image responses under continuous naturalistic stimuli. By integrating statistical methodology with modern computational frameworks, this work provides new tools for understanding both the higher-order connectivity and the dynamic effect patterns in the human brain.

APPENDIX A

Appendix of Chapter I

A.1 Proof of Technical Lemmas

Under Condition 2.1, we study the statistical rate of the sample moment estimator (3.2.3) under the Frobenius norm. To see this, we start with the upper bound of the max norm

$$\max_{\{j,k,s,t\}\in\{1,\ldots,p\}^4} \left|\widehat{\Theta}_{jk,st} - \Theta_{jk,st}\right|,$$

which consists of the second-order and fourth-order moments in the Lemma A.1.1 and Lemma A.1.2, respectively.

Lemma A.1.1. Under Condition 2.1, we have

$$\max_{\{j,k,s,t\}\in\{1,\ldots,p\}^4} \left|\frac{1}{n}\sum_{i=1}^{n} X_{ij}X_{ik} \cdot \frac{1}{n}\sum_{i=1}^{n} X_{is}X_{it} - \sigma_{jk}\sigma_{st}\right| \leq C_1\sqrt{\frac{\log(np)}{n}} + C_2\frac{\log(np)}{n},$$

for some constants C_1 and C_2, with probability at least $1 - n^{-1}$.

A.1.1 Proof of Lemma A.1.1

Proof. We start the proof by decomposing the quantity $|n^{-1}\sum_{i=1}^{n} X_{ij}X_{ik} \cdot n^{-1}\sum_{i=1}^{n} X_{is}X_{it} - \sigma_{jk}\sigma_{st}|$ by adding and subtracting terms:

$$\begin{aligned}
&\left|\frac{1}{n}\sum_{i=1}^{n} X_{ij}X_{ik} \cdot \frac{1}{n}\sum_{i=1}^{n} X_{is}X_{it} - \sigma_{jk}\sigma_{st}\right| \\
&= \left|\left(\frac{1}{n}\sum_{i=1}^{n} X_{ij}X_{ik} - \sigma_{jk}\right) \cdot \frac{1}{n}\sum_{i=1}^{n} X_{is}X_{it} + \sigma_{jk}\left(\frac{1}{n}\sum_{i=1}^{n} X_{is}X_{it} - \sigma_{st}\right)\right| \\
&\leq \underbrace{\left|\left(\frac{1}{n}\sum_{i=1}^{n} X_{ij}X_{ik} - \sigma_{jk}\right) \cdot \left(\frac{1}{n}\sum_{i=1}^{n} X_{is}X_{it} - \sigma_{st}\right)\right|}_{I} + \underbrace{\left|\left(\frac{1}{n}\sum_{i=1}^{n} X_{ij}X_{ik} - \sigma_{jk}\right)\sigma_{st}\right|}_{II}
\end{aligned}$$

$$+ \underbrace{\left| \sigma_{jk} \left(\frac{1}{n} \sum_{i=1}^{n} X_{is} X_{it} - \sigma_{st} \right) \right|}_{III}.$$

It suffices to show that I, II, and III are upper bounded by $C_1\sqrt{n^{-1}\log(np)} + C_2 n^{-1}\log(np)$ for some constants C_1, C_2.

By an application of Lemma 1 from Ravikumar et al. (2011), for some constant C, we have

$$II \leq C\sqrt{\frac{\log(n)}{n}},$$

with probability at least $1 - n^{-1}$. By union bound, we further obtain

$$\max_{\{j,k,s,t\}\in\{1,\ldots,p\}^4} II \leq C_1\sqrt{\frac{\log(np)}{n}},$$

with probability at least $1 - n^{-1}$ for some constant C_1. We obtain the same result for III.

For I, note that

$$\max_{\{j,k,s,t\}\in\{1,\ldots,p\}^4} I \leq \max_{\{j,k,s,t\}\in\{1,\ldots,p\}^4} \left(\frac{II}{|\sigma_{st}|} \frac{III}{|\sigma_{jk}|} \right).$$

By union bound, for some constant C', we have

$$\begin{aligned}
&\mathbb{P}\left(\max_{\{j,k,s,t\}\in\{1,\ldots,p\}^4} \left(\frac{II}{|\sigma_{st}|} \frac{III}{|\sigma_{jk}|} \right) \geq C'^2 \frac{\log(np)}{n} \right) \\
&\leq \mathbb{P}\left(\left\{ \max_{\{j,k,s,t\}\in\{1,\ldots,p\}^4} \frac{II}{|\sigma_{st}|} \geq C'\sqrt{\frac{\log(np)}{n}} \right\} \right. \\
&\qquad \left. \bigcup \left\{ \max_{\{j,k,s,t\}\in\{1,\ldots,p\}^4} \frac{III}{|\sigma_{jk}|} \geq C'\sqrt{\frac{\log(np)}{n}} \right\} \right) \\
&\leq \mathbb{P}\left(\max_{\{j,k,s,t\}\in\{1,\ldots,p\}^4} \frac{II}{|\sigma_{st}|} \geq C'\sqrt{\frac{\log(np)}{n}} \right) \\
&\qquad + \mathbb{P}\left(\max_{\{j,k,s,t\}\in\{1,\ldots,p\}^4} \frac{III}{|\sigma_{jk}|} \geq C'\sqrt{\frac{\log(np)}{n}} \right) \\
&\leq \frac{2}{n}.
\end{aligned}$$

Therefore, for some constant C_2,

$$\max_{\{j,k,s,t\}\in\{1,\ldots,p\}^4} I \leq C_2 \frac{\log(np)}{n}$$

with probability at least $1 - 2n^{-1}$.

Combining the aforementioned upper bounds on I, II, and III, we obtain the desired results. □

It remains to find the concentration inequality for the fourth-order moment part.

Lemma A.1.2. Under Condition 2.1, we have

$$\max_{\{j,k,s,t\}\in\{1,\cdots,p\}^4} \left| \frac{1}{n}(\sum_{i=1}^{n} X_{ij}X_{ik}X_{is}X_{it} - \mathbb{E}[X_{ij}X_{ik}X_{is}X_{it}]) \right| \leq C_3\sqrt{\frac{\log^4(np^4)}{n}},$$

for some constant C_3 with probability at least $1 - n^{-1}$.

A.1.2 Proof of Lemma A.1.2

Proof. We aim to find the bound u, so that

$$\mathbb{P}\left(\max_{\{j,k,s,t\}\in\{1,\ldots,p\}^4} \left| \frac{1}{n}(\sum_{i=1}^{n} X_{ij}X_{ik}X_{is}X_{it} - \mathbb{E}[X_{ij}X_{ik}X_{is}X_{it}]) \right| > u \right) < \frac{1}{n}.$$

We start with union bound,

$$\begin{aligned}
&\mathbb{P}(\max_{\{j,k,s,t\}\in\{1,\ldots,p\}^4} |\frac{1}{n}(\sum_{i=1}^{n} X_{ij}X_{ik}X_{is}X_{it} - \mathbb{E}[X_{ij}X_{ik}X_{is}X_{it}])| > u) \\
&= \mathbb{P}(\bigcup_{\{j,k,s,t\}\in\{1,\ldots,p\}^4} |\frac{1}{n}(\sum_{i=1}^{n} X_{ij}X_{ik}X_{is}X_{it} - \mathbb{E}[X_{ij}X_{ik}X_{is}X_{it}])| > u) \\
&\leq \sum_{\{j,k,s,t\}\in\{1,\ldots,p\}^4} \mathbb{P}(|(\sum_{i=1}^{n} X_{ij}X_{ik}X_{is}X_{it} - \mathbb{E}[X_{ij}X_{ik}X_{is}X_{it}])| > nu) \\
&\leq \sum_{\{j,k,s,t\}\in\{1,\ldots,p\}^4} (nu)^{-v}\mathbb{E}[(\sum_{i=1}^{n} X_{ij}X_{ik}X_{is}X_{it} - \mathbb{E}[X_{ij}X_{ik}X_{is}X_{it}])^v],
\end{aligned}$$

where the last inequality follows from an application of Markov inequality.

It suffices to obtain the upper bound of $\mathbb{E}[(\sum_{i=1}^{n} X_{ij}X_{ik}X_{is}X_{it} - \mathbb{E}[X_{ij}X_{ik}X_{is}X_{it}])^v]$. Note that by the product of sum, we can rewrite the expectation as

$$\begin{aligned}
&\mathbb{E}[(\sum_{i=1}^{n} X_{ij}X_{ik}X_{is}X_{it} - \mathbb{E}[X_{ij}X_{ik}X_{is}X_{it}])^v] \\
&= \sum_{i_1,\ldots,i_n\in[n]} \mathbb{E}[\prod_{q=1}^{v}(X_{ij_q}X_{ik_q}X_{is_q}X_{it_q} - \mathbb{E}[X_{ij_q}X_{ik_q}X_{is_q}X_{it_q}])] \\
&= \sum_{v_1+\ldots+v_n=v/2} \prod_{i=1}^{n} \mathbb{E}(X_{ij_q}X_{ik_q}X_{is_q}X_{it_q} - \mathbb{E}[X_{ij_q}X_{ik_q}X_{is_q}X_{it_q}])^{2v_i}
\end{aligned}$$

$$
\begin{aligned}
&\leq \sum_{v_1+\ldots+v_n=v/2} \prod_{i=1}^{n} \mathbb{E}(|X_{ij_q}X_{ik_q}X_{is_q}X_{it_q}| + |\mathbb{E}[X_{ij_q}X_{ik_q}X_{is_q}X_{it_q}]|)^{2v_i} \\
&= \sum_{v_1+\ldots+v_n=v/2} \prod_{i=1}^{n} \mathbb{E}(\sum_{m=0}^{2v_i} \binom{2v_i}{m} |X_{ij_q}X_{ik_q}X_{is_q}X_{it_q}|^{2v_i-m} |\mathbb{E}[X_{ij_q}X_{ik_q}X_{is_q}X_{it_q}]|^m),
\end{aligned}
$$

where the inequality is obtained by the triangle inequality and the last equality is obtained by rewriting the product of sum.

By Cauchy-Schwarz inequality, we have

$$
\begin{aligned}
|\mathbb{E}[X_{ij_q}X_{ik_q}X_{is_q}X_{it_q}]|^m &\leq (\mathbb{E}[X_{ij_q}^2 X_{ik_q}^2])^{m/2} (\mathbb{E}[X_{is_q}^2 X_{it_q}^2])^{m/2} \\
&\leq (\mathbb{E}[X_{ij_q}^4]\mathbb{E}[X_{ik_q}^4]\mathbb{E}[X_{is_q}^4]\mathbb{E}[X_{it_q}^4])^{m/4} (\mathbb{E}[X_{is_q}^2 X_{it_q}^2])^{m/2} \quad \text{(A.1.1)} \\
&\leq K^{4m} 4^{2m},
\end{aligned}
$$

for some constant K, where the last inequality is obtained by Condition 2.1.

Applying the same process for $\mathbb{E}(|X_{ij_q}X_{ik_q}X_{is_q}X_{it_q}|^{2p_i-m})$, we have

$$
\begin{aligned}
&\sum_{v_1+\ldots+v_n=v/2} \prod_{i=1}^{n} \mathbb{E}(\sum_{m=0}^{2v_i} \binom{2v_i}{m} |X_{ij_q}X_{ik_q}X_{is_q}X_{it_q}|^{2v_i-m} |\mathbb{E}[X_{ij_q}X_{ik_q}X_{is_q}X_{it_q}]|^m) \\
&\leq \sum_{v_1+\ldots+v_n=v/2} \prod_{i=1}^{n} \sum_{m=0}^{2v_i} \binom{2v_i}{m} (\mathbb{E}[X_{ij_q}^{4v_i-2m} X_{ik_q}^{4v_i-2m}])^{1/2} (\mathbb{E}[X_{is_q}^{4v_i-2m} X_{it_q}^{4v_i-2m}])^{1/2} K^{4m} 4^{2m} \\
&\leq \sum_{v_1+\ldots+v_n=v/2} \prod_{i=1}^{n} \sum_{m=0}^{2v_i} \binom{2v_i}{m} (\mathbb{E}[X_{ij_q}^{8v_i-4m}]\mathbb{E}[X_{ik_q}^{8v_i-4m}]\mathbb{E}[X_{is_q}^{8v_i-4m}]\mathbb{E}[X_{it_q}^{8v_i-4m}])^{1/4} K^{4m} 4^{2m} \\
&\leq \sum_{v_1+\ldots+v_n=v/2} \prod_{i=1}^{n} \sum_{m=0}^{2v_i} \binom{2v_i}{m} [K^{(8v_i-4m)4} (8v_i-4m)^{(4v_i-2m)4}]^{1/4} K^{4m} 4^{2m} \\
&= \sum_{v_1+\ldots+v_n=v/2} \prod_{i=1}^{n} \sum_{m=0}^{2v_i} \binom{2v_i}{m} K^{8v_i} 4^{4v_i} (2v_i-m)^{4v_i-2m} \\
&\leq (Cv^4 n)^{v/2},
\end{aligned}
$$

for some constant C, where the last inequality holds by an application of Lemma A.2.1.

Choose $v = \log(np^4)$, $u = C_3\sqrt{\log^4(np^4)/n}$ and we obtain the desired results. □

Lemma A.1.3. Under Condition 2.1, we have

$$
\max_{\{j,k,s,t\}\in\{1,\ldots,p\}^4} \left|\widehat{\Theta}_{jk,st} - \Theta_{jk,st}\right| \leq C_1\sqrt{\frac{\log(np)}{n}} + C_2\frac{\log(np)}{n} + C_3\sqrt{\frac{\log^4(np^4)}{n}},
$$

for some constants C_1, C_2, C_3, with probability at least $1 - n^{-1}$.

The proof can be found in Appendix A.1.3.

A.1.3 Proof of Lemma A.1.3

Proof. By triangle inequality, we decompose the quantity:

$$\begin{aligned}&\left|\widehat{\Theta}_{jk,st} - \Theta_{jk,st}\right| \\ &\leq \left|\frac{1}{n}\sum_{i=1}^{n} X_{ij}X_{ik}X_{is}X_{it} - \mathbb{E}[X_jX_kX_sX_t]\right| + \left|\frac{1}{n}\sum_{i=1}^{n} X_{ij}X_{ik} \cdot \frac{1}{n}\sum_{i=1}^{n} X_{is}X_{it} - \sigma_{jk}\sigma_{st}\right|.\end{aligned}$$

Applying Lemma A.1.1 and A.1.2, we get the desired result. □

A.1.4 Proof of Theorem 2.2.1

Proof. By Lemma A.1.3, we have

$$|\widehat{\Theta}_{jk,st} - \Theta_{jk,st}| \leq \max_{\{j,k,s,t\}} |\widehat{\Theta}_{jk,st} - \Theta_{jk,st}| \leq C'\sqrt{\frac{\log^4(np^4)}{n}},$$

for some constant C' with probability at least $1 - n^{-1}$.

Then we can write

$$\begin{aligned}\|\widehat{\mathbf{\Theta}} - \mathbf{\Theta}\|_F^2 &= \sum_{\{j,k,s,t\}} (\widehat{\Theta}_{jk,st} - \Theta_{jk,st})^2 \\ &\leq \sum_{\{j,k,s,t\}} \max_{\{j,k,s,t\}} (\widehat{\Theta}_{jk,st} - \Theta_{jk,st})^2 \\ &\leq Cp^4\frac{\log^4(np^4)}{n},\end{aligned}$$

for some constant C with probability at least $1 - n^{-1}$. □

A.1.5 Proof of Theorem 2.2.3

Proof. We start the proof by decomposing $\|T_t(\widehat{\mathbf{\Theta}}) - \mathbf{\Theta}\|_F^2$ into

$$\sum_{\{j,k,s,t\}} (T_t(\widehat{\Theta}_{jk,st}) - \Theta_{jk,st})^2 \mathbb{1}(|\Theta_{jk,st}| \neq 0) + \sum_{\{j,k,s,t\}} (T_t(\widehat{\Theta}_{jk,st}) - \mathbf{\Theta}_{jk,st})^2 \mathbb{1}(|\Theta_{jk,st}| = 0).$$

By Lemma A.1.3, we have

$$|\widehat{\Theta}_{jk,st} - \Theta_{jk,st}| \leq \max_{\{j,k,s,t\}} |\widehat{\Theta}_{jk,st} - \Theta_{jk,st}| \leq C\sqrt{\frac{\log^4(np^4)}{n}},$$

for some constants C with probability at least $1 - n^{-1}$. Take $t = C\sqrt{\log^4(np^4)/n}$. When $\Theta_{jk,st} = 0$, $|\widehat{\Theta}_{jk,st}| \leq t$ with probability at least $1 - n^{-1}$.

Then $\|T_t(\widehat{\Theta}_{jk,st}) - \Theta_{jk,st}\|_F^2$ becomes

$$\begin{aligned}
&\sum_{\{j,k,s,t\}} (\widehat{\Theta}_{jk,st} - \Theta_{jk,st})^2 \mathbb{1}(|\Theta_{jk,st}| \neq 0) + \sum_{\{j,k,s,t\}} (T_t(\widehat{\Theta}_{jk,st}))^2 \mathbb{1}(|\widehat{\Theta}_{jk,st}| \leq t) \\
&= \sum_{\{j,k,s,t\}} (\widehat{\Theta}_{jk,st} - \Theta_{jk,st})^2 \mathbb{1}(|\Theta_{jk,st}| \neq 0) \\
&\leq \max_{\{j,k,s,t\}} (\widehat{\Theta}_{jk,st} - \Theta_{jk,st})^2 \sum_{\{j,k,s,t\}} \mathbb{1}(|\Theta_{jk,st}| \neq 0) \leq t^2 m,
\end{aligned}$$

with probability at least $1 - n^{-1}$. □

A.1.6 Proof of Theorem 2.2.2

Proof. By law of large number and Slutsky's theorem, the sample estimate $\widehat{\Lambda}_{jk,st} \xrightarrow{p} \Lambda_{jk,st}$. Based on the asymptotic normality of $\widehat{\Theta}_{jk,st}$, we further obtain $\gamma_{jk,st} \xrightarrow{d} \mathcal{N}(0,1)$ as $n \to \infty$.

For any $\alpha \in (0,1)$, the event $\{p_{(l)} \leq \alpha\}$ is equivalent to the event $\{|\gamma_{jk,st}| \geq z_{\alpha/2}\}$. As $\gamma_{jk,st} \xrightarrow{d} \mathcal{N}(0,1)$, $\mathbb{P}(|\gamma_{jk,st}| \geq z_{\alpha/2}) \longrightarrow \alpha$. We obtain

$$\mathbb{P}(p_{(l)} \leq \alpha) = \mathbb{P}(|\gamma_{jk,st}| \geq z_{\alpha/2}) = \alpha + o(1). \tag{A.1.2}$$

For a set of N hypotheses, denote $\mathcal{H}_0$ as the set of true null hypotheses and R as the number of rejections. By the procedure of FDR control, we have

$$R = \sum_{l=1}^{N} \mathbb{1}\left(p_l \leq \frac{l}{NL}\alpha\right),$$

where $L = \sum_{l=1}^{N} 1/l$. Following the proof of Theorem 4.2 in Liu et al. (2022a), we have

$$\frac{1}{R} = \sum_{r=R}^{\infty} \frac{1}{r(r+1)} = \sum_{r=1}^{\infty} \frac{\mathbb{1}(r \geq R)}{r(r+1)},$$

and

$$\begin{aligned}
\text{FDR} &= \sum_{l \in \mathcal{H}_0} \mathbb{E}\left[\frac{\mathbb{1}\left(p_l \leq \frac{R}{NL}\alpha\right)\mathbb{1}(R>0)}{R}\right] \\
&= \sum_{r=1}^{\infty} \frac{\mathbb{1}(r \geq R)}{r(r+1)} \sum_{l \in \mathcal{H}_0} \mathbb{E}\left[\mathbb{1}\left(p_l \leq \frac{R}{NL}\alpha\right)\mathbb{1}(R>0)\right] \\
&\leq \sum_{r=1}^{\infty} \frac{1}{r(r+1)} \sum_{l \in \mathcal{H}_0} \mathbb{P}\left(p_l \leq \frac{\min(r,N)}{NL}\alpha\right).
\end{aligned}$$

By (A.1.2), we further obtain

$$\begin{aligned}
\text{FDR} &\leq \sum_{r=1}^{\infty} \frac{1}{r(r+1)} \sum_{l \in \mathcal{H}_0} \left(\frac{\min(r, N)}{NL} \alpha + o(1) \right) \\
&= \sum_{r=1}^{N} \frac{1}{r+1} \sum_{l \in \mathcal{H}_0} \frac{1}{NL} \alpha + \sum_{r=N+1}^{\infty} \frac{1}{r(r+1)} \sum_{l \in \mathcal{H}_0} \frac{1}{L} \alpha + o(1) \\
&= \sum_{l \in \mathcal{H}_0} \frac{\alpha}{NL} \left(\sum_{r=1}^{N} \frac{1}{r+1} + \sum_{r=N+1}^{\infty} \frac{N}{r(r+1)} \right) + o(1) \\
&= \sum_{l \in \mathcal{H}_0} \frac{\alpha}{NL} \left(\sum_{r=1}^{N} \frac{1}{r+1} + \frac{N}{N+1} \right) + o(1) \\
&= \sum_{l \in \mathcal{H}_0} \frac{\alpha}{NL} \cdot \sum_{r=1}^{N} \frac{1}{r} + o(1) \\
&= \sum_{l \in \mathcal{H}_0} \frac{\alpha}{N} + o(1) \\
&\leq \alpha + o(1).
\end{aligned}$$

□

A.2 Some Useful Technical Lemmas

In this section, we will outline some technical lemmas that will be used to prove our main results.

Lemma A.2.1. For some positive even constant v and some constant $K > 0$,

$$\sum_{v_1 + \cdots + v_n = v/2} \prod_{i=1}^{n} \sum_{m=0}^{2v_i} \binom{2v_i}{m} K^{8v_i} 4^{4v_i} (2v_i - m)^{4v_i - 2m} \leq (Cv^4 n)^{v/2},$$

for some constant C.

Proof. First, note that

$$\binom{n}{k} \leq \left(\frac{ne}{k} \right)^k,$$

then we have

$$\begin{aligned}
&\sum_{v_1 + \cdots + v_n = v/2} \prod_{i=1}^{n} \sum_{m=0}^{2v_i} \binom{2v_i}{m} K^{8v_i} 4^{4v_i} (2v_i - m)^{4v_i - 2m} \\
&= \sum_{v_1 + \cdots + v_n = v/2} \prod_{i=1}^{n} \sum_{m=0}^{2v_i} \binom{2v_i}{2v_i - m} K^{8v_i} 4^{4v_i} (2v_i - m)^{4v_i - 2m} \\
&\leq \sum_{v_1 + \ldots + v_n = v/2} \prod_{i=1}^{n} K^{8v_i} 4^{4v_i} \sum_{m=0}^{2v_i} (2v_i e)^{2v_i - m} (2v_i - m)^{2v_i - m}
\end{aligned}$$

$$\leq \sum_{v_1+\ldots+v_n=v/2} \prod_{i=1}^{n} K^{8v_i} 4^{4v_i} \sum_{m=0}^{2v_i} (2v_i e)^{2v_i-m} (2v_i)^{2v_i-m}. \tag{A.2.1}$$

By the sum of a geometric series, we can further simplify (A.2.1) to

$$\sum_{v_1+\ldots+v_n=v/2} \prod_{i=1}^{n} K^{8v_i} 4^{4v_i} \frac{(4v_i^2 e)^{2v_i+1}-1}{(4v_i^2 e)-1} \leq \sum_{v_1+\ldots+v_n=v/2} \prod_{i=1}^{n} K^{8v_i} 4^{4v_i} \frac{(4v_i^2 e)^{2v_i+1}}{(4v_i^2 e)-1}. \tag{A.2.2}$$

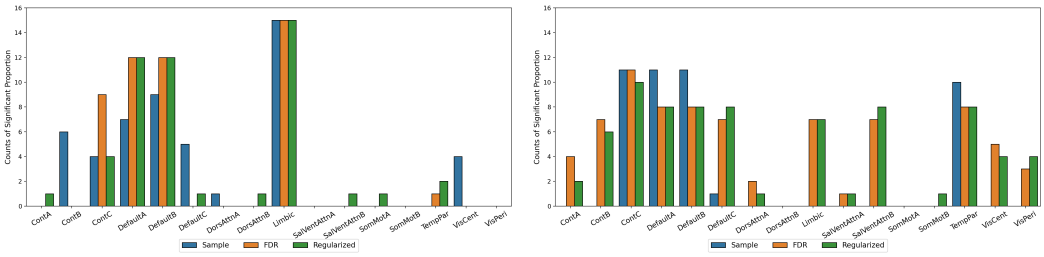
Since $v_i > 1$, $4v_i^2 e > 2$ and $4v_i^2 e - 1 > 4v_i^2 e/2$, (A.2.2) can be upper bounded by

$$\begin{aligned}
&\sum_{v_1+\ldots+v_n=v/2} K^{4v} 4^{2v} \prod_{i=1}^{n} (4v_i^2 e)^{2v_i+1} \frac{2}{4v_i^2 e} \\
&= \sum_{v_1+\ldots+v_n=v/2} K^{4v} 4^{2v} 2^n (v^2 e)^v \\
&= \binom{v/2+n-1}{v/2} (16K^4 e)^v 2^n v^{2v} \\
&\leq (\frac{(v/2+n-1)e}{v/2})^{v/2} (256K^8 e^2)^{v/2} (2^{2n/v})^{v/2} (v^4)^{v/2} \leq (Cv^4 n)^{v/2},
\end{aligned}$$

for some constant C, where the first inequality is obtained by an application of the binomial inequality.

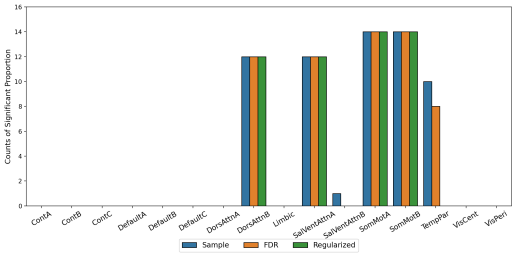
□

A.3 Additional Numerical Results of ceFC

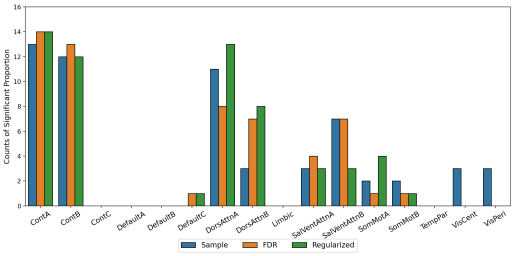


(a) Community 1

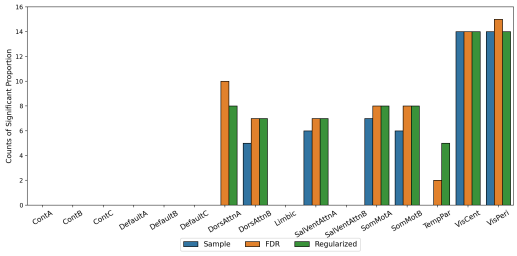
(b) Community 2



(c) Community 3

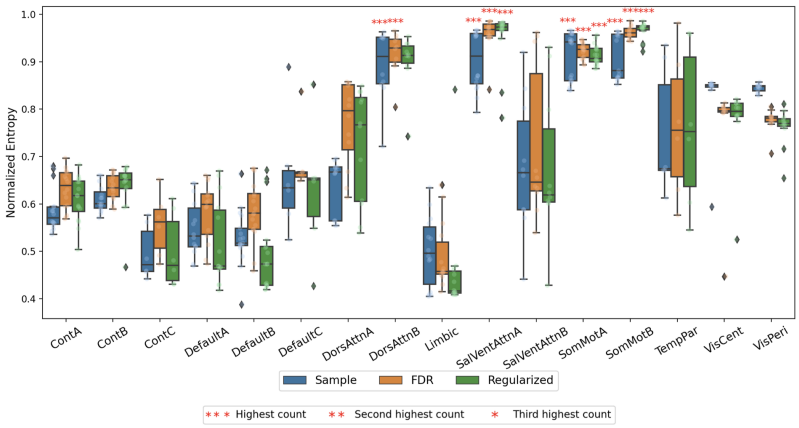


(d) Community 4



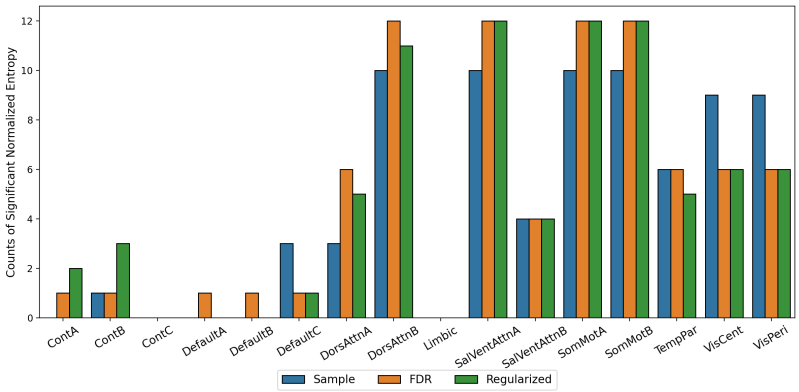
(e) Community 5

Figure A.1: Counts of significant outcomes from pairwise comparisons of edge proportions among brain systems. The top three brain systems, identified by their top three highest counts within each method, are highlighted in Figure 4.



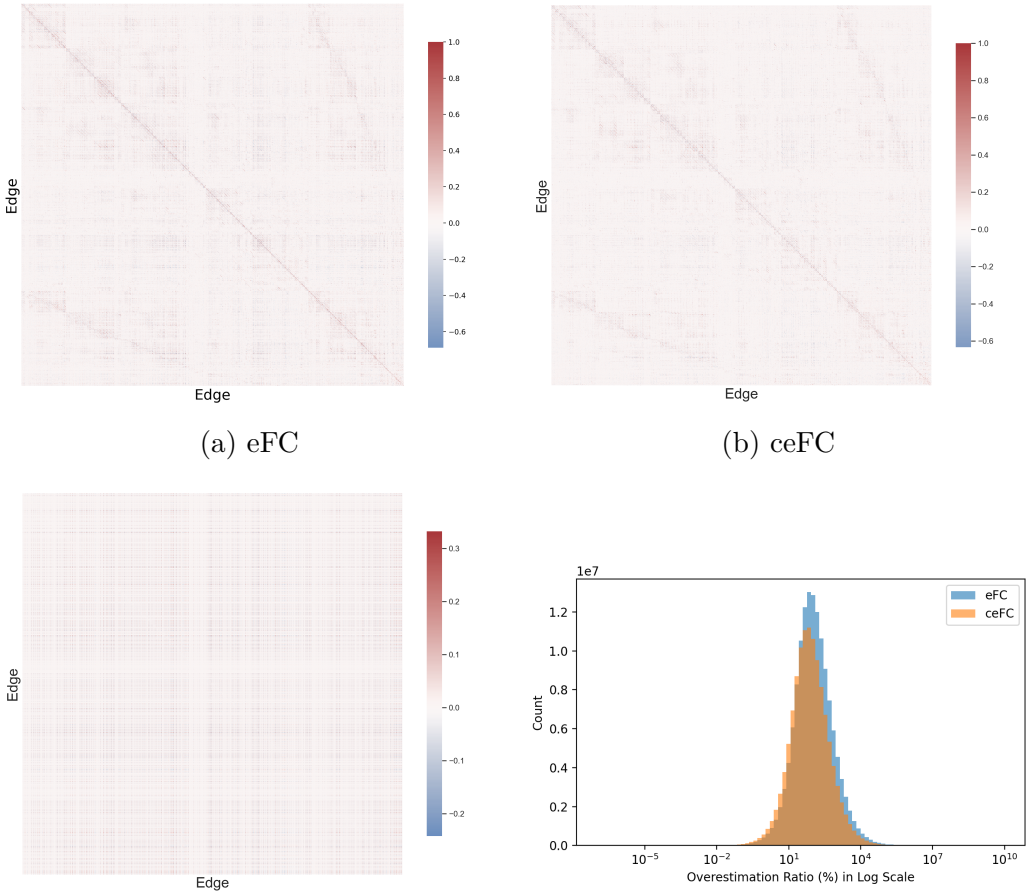
(a)

Figure A.2: Normalized entropy of edge proportions in Figure 3 over all communities. The plot highlights the top three brain systems that demonstrate a significantly greater normalized entropy of edges in pairwise comparisons with other systems. $***$ denotes the system with the highest count of significant outcomes, $**$ represents the second highest, and $*$ signifies the third highest.



(a)

Figure A.3: Counts of significant outcomes from pairwise comparisons of normalized entropy among brain systems. The top three brain systems, identified by their top three highest counts within each method, are highlighted in Figure A.2.



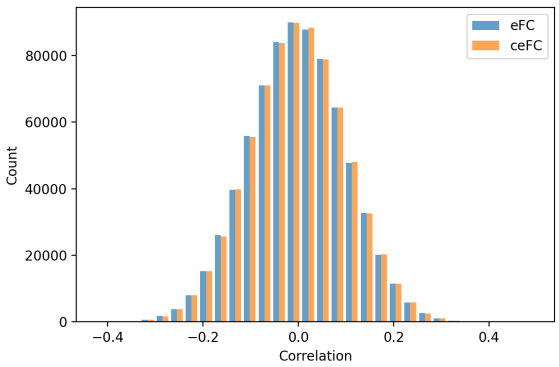
(a) eFC

(b) ceFC

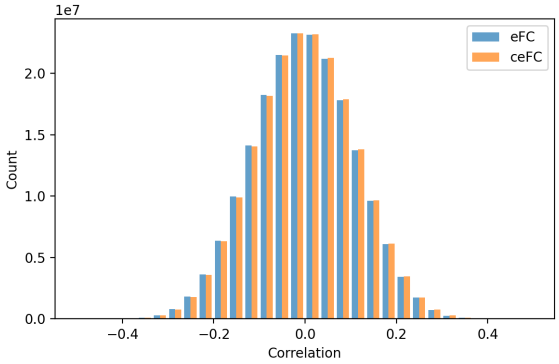
(c) eFC - ceFC

(d) Overestimation Ratio

Figure A.4: Sample moment estimation of eFC, ceFC, and eFC − ceFC in (a), (b), and (c), respectively. Plot (d) shows the distribution of the overestimation ratio (%) on a log scale.



(a) $p = 50$



(b) $p = 200$

Figure A.5: Correlation of each method from two halves of the simulation as split-half reliability. Plots (a) and (b) show the count of edge pairs for corresponding correlation values with $p = 50$ and $p = 200$ parcels, respectively.

APPENDIX B

Appendix of Chapter II

B.1 Comparison between PLS and Linear Regression with Orthogonal Covariates

In Algorithm 2, we show that for PLS with orthogonal covariates, the resulting coefficient $\boldsymbol{B}_{PLS}$ can be obtained by

$$\begin{aligned}
\boldsymbol{B}_{PLS} &= \boldsymbol{R}\boldsymbol{T}^\top \boldsymbol{Y} \\
&= \frac{\boldsymbol{X}^\top \boldsymbol{Y}\boldsymbol{Y}^\top \boldsymbol{X}\boldsymbol{X}^\top \boldsymbol{Y}}{\boldsymbol{Y}^\top \boldsymbol{X}\boldsymbol{K}\boldsymbol{X}^\top \boldsymbol{Y}} \\
&= \frac{\boldsymbol{X}^\top \boldsymbol{Y}\boldsymbol{Y}^\top \boldsymbol{X}\boldsymbol{K}}{\boldsymbol{Y}^\top \boldsymbol{X}\boldsymbol{K}\boldsymbol{X}^\top \boldsymbol{Y}} \left(\boldsymbol{K}^{-1}\boldsymbol{X}^\top \boldsymbol{Y}\right) \\
&= \frac{\boldsymbol{X}^\top \boldsymbol{Y}\boldsymbol{Y}^\top \boldsymbol{X}\boldsymbol{K}}{\boldsymbol{Y}^\top \boldsymbol{X}\boldsymbol{K}\boldsymbol{X}^\top \boldsymbol{Y}} \boldsymbol{B}_{OLS},
\end{aligned}$$

where $\boldsymbol{B}_{OLS} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{Y}$ is the linear regression coefficient via ordinary least squares.

When $\boldsymbol{X}$ is orthonormal, i.e., $\boldsymbol{X}^\top \boldsymbol{X} = \boldsymbol{K} = \boldsymbol{I}$, we have

$$\begin{aligned}
\frac{\boldsymbol{X}^\top \boldsymbol{Y}\boldsymbol{Y}^\top \boldsymbol{X}\boldsymbol{K}}{\boldsymbol{Y}^\top \boldsymbol{X}\boldsymbol{K}\boldsymbol{X}^\top \boldsymbol{Y}} &= \frac{\boldsymbol{X}^\top \boldsymbol{Y}\boldsymbol{Y}^\top \boldsymbol{X}}{\boldsymbol{Y}^\top \boldsymbol{X}\boldsymbol{X}^\top \boldsymbol{Y}} \\
&= \frac{\boldsymbol{V}\boldsymbol{V}^\top}{\boldsymbol{V}^\top \boldsymbol{V}} \\
&= \boldsymbol{V}(\boldsymbol{V}^\top \boldsymbol{V})^{-1}\boldsymbol{V}^\top \\
&= \boldsymbol{P},
\end{aligned}$$

where $\boldsymbol{V} = \boldsymbol{X}^\top \boldsymbol{Y}$ and $\boldsymbol{P}$ is a projection matrix. Then, $\boldsymbol{B}_{OLS} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{Y} = \boldsymbol{V}$. Therefore, we have $\boldsymbol{P}\boldsymbol{V} = \boldsymbol{V}$. That is,

$$\boldsymbol{B}_{PLS} = \frac{\boldsymbol{X}^\top \boldsymbol{Y}\boldsymbol{Y}^\top \boldsymbol{X}\boldsymbol{K}}{\boldsymbol{Y}^\top \boldsymbol{X}\boldsymbol{K}\boldsymbol{X}^\top \boldsymbol{Y}} \boldsymbol{B}_{OLS} = \boldsymbol{B}_{OLS}.$$

In particular, assume $\boldsymbol{K} = \boldsymbol{X}^\top \boldsymbol{X} = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_p \end{bmatrix}$. Then

$$\boldsymbol{B}_{OLS} = \begin{bmatrix} \frac{X_1^\top \boldsymbol{Y}}{d_1} \\ \vdots \\ \frac{X_p^\top \boldsymbol{Y}}{d_p} \end{bmatrix},$$

$$\boldsymbol{X}^\top \boldsymbol{Y}\boldsymbol{Y}^\top \boldsymbol{X}\boldsymbol{K} = \begin{bmatrix} d_1 X_1^\top \boldsymbol{Y}\boldsymbol{Y}^\top X_1 & \cdots & d_p X_1^\top \boldsymbol{Y}\boldsymbol{Y}^\top X_p \\ \vdots & \cdots & \vdots \\ d_1 X_p^\top \boldsymbol{Y}\boldsymbol{Y}^\top X_1 & \cdots & d_p X_p^\top \boldsymbol{Y}\boldsymbol{Y}^\top X_p \end{bmatrix},$$

$$\boldsymbol{Y}^\top \boldsymbol{X}\boldsymbol{K}\boldsymbol{X}^\top \boldsymbol{Y} = \sum_{j=1}^{p} d_j (\boldsymbol{Y}^\top X_j X_j^\top \boldsymbol{Y}) = \sum_{j=1}^{p} d_j (X_j^\top \boldsymbol{Y})^2.$$

Plug in $X_j^\top \boldsymbol{Y} = \boldsymbol{Y}^\top X_j = d_j B_{j_{(OLS)}}$ and $\boldsymbol{Y}^\top X_j X_j^\top \boldsymbol{Y} = d_j^2 B_{j_{(OLS)}}^2$, we can see

$$\begin{aligned} \frac{\boldsymbol{X}^\top \boldsymbol{Y}\boldsymbol{Y}^\top \boldsymbol{X}\boldsymbol{K}}{\boldsymbol{Y}^\top \boldsymbol{X}\boldsymbol{K}\boldsymbol{X}^\top \boldsymbol{Y}} \boldsymbol{B}_{OLS} &= \begin{bmatrix} \frac{X_1^\top \boldsymbol{Y}\boldsymbol{Y}^\top X_1 \cdot (X_1^\top \boldsymbol{Y})}{\sum_{j=1}^p d_j (X_j^\top \boldsymbol{Y})^2} + \cdots + \frac{X_1^\top \boldsymbol{Y}\boldsymbol{Y}^\top X_p \cdot (X_p^\top \boldsymbol{Y})}{\sum_{j=1}^p d_j (X_j^\top \boldsymbol{Y})^2} \\ \vdots \\ \frac{X_p^\top \boldsymbol{Y}\boldsymbol{Y}^\top X_1 \cdot (X_1^\top \boldsymbol{Y})}{\sum_{j=1}^p d_j (X_j^\top \boldsymbol{Y})^2} + \cdots + \frac{X_p^\top \boldsymbol{Y}\boldsymbol{Y}^\top X_p \cdot (X_p^\top \boldsymbol{Y})}{\sum_{j=1}^p d_j (X_j^\top \boldsymbol{Y})^2} \end{bmatrix} \\ &= \begin{bmatrix} \frac{X_1^\top \boldsymbol{Y} \cdot d_1^2 B_{1_{(OLS)}}^2}{\sum_{j=1}^p d_j^3 B_{j_{(OLS)}}^2} + \cdots + \frac{X_1^\top \boldsymbol{Y} \cdot d_p^2 B_{p_{(OLS)}}^2}{\sum_{j=1}^p d_j^3 B_{j_{(OLS)}}^2} \\ \vdots \\ \frac{X_p^\top \boldsymbol{Y} \cdot d_1^2 B_{1_{(OLS)}}^2}{\sum_{j=1}^p d_j^3 B_{j_{(OLS)}}^2} + \cdots + \frac{X_p^\top \boldsymbol{Y} \cdot d_p^2 B_{p_{(OLS)}}^2}{\sum_{j=1}^p d_j^3 B_{j_{(OLS)}}^2} \end{bmatrix} \\ &= \begin{bmatrix} X_1^\top \boldsymbol{Y} \cdot \frac{\sum_{i=1}^p d_i^2 B_{i_{(OLS)}}^2}{\sum_{j=1}^p d_j^3 B_{j_{(OLS)}}^2} \\ \vdots \\ X_p^\top \boldsymbol{Y} \cdot \frac{\sum_{i=1}^p d_i^2 B_{i_{(OLS)}}^2}{\sum_{j=1}^p d_j^3 B_{j_{(OLS)}}^2}. \end{bmatrix} \end{aligned}$$

When all d_js are equal, we have

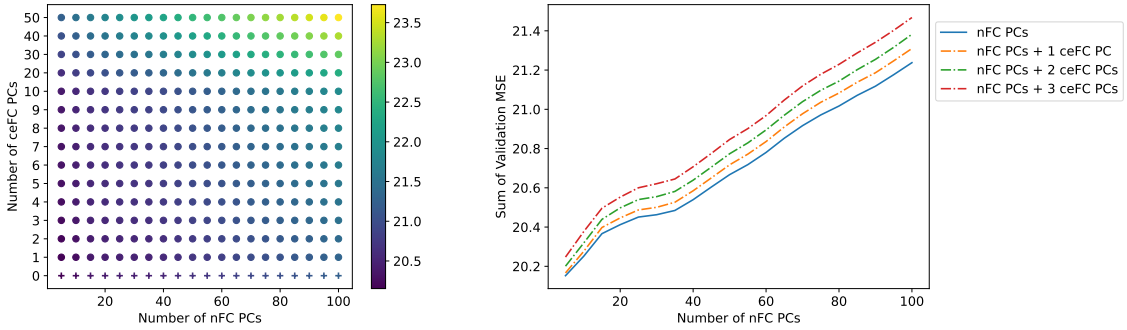
$$\boldsymbol{B}_{PLS} = \frac{\boldsymbol{X}^\top \boldsymbol{Y}\boldsymbol{Y}^\top \boldsymbol{X}\boldsymbol{K}}{\boldsymbol{Y}^\top \boldsymbol{X}\boldsymbol{K}\boldsymbol{X}^\top \boldsymbol{Y}} \boldsymbol{B}_{OLS} = \begin{bmatrix} \frac{X_1^\top \boldsymbol{Y}}{d} \\ \vdots \\ \frac{X_p^\top \boldsymbol{Y}}{d} \end{bmatrix} = \boldsymbol{B}_{OLS}.$$

Plug in the coefficients for prediction, we obtain

$$\begin{aligned}\widehat{\boldsymbol{Y}} = \boldsymbol{X}\boldsymbol{B}_{PLS} &= \frac{\sum_{i=1}^{p} d_i^2 B_{i_{(OLS)}}^2}{\sum_{j=1}^{p} d_j^3 B_{j_{(OLS)}}^2} X_1^\top \boldsymbol{Y} X_1 + \cdots + \frac{\sum_{i=1}^{p} d_i^2 B_{i_{(OLS)}}^2}{\sum_{j=1}^{p} d_j^3 B_{j_{(OLS)}}^2} \cdot X_p^\top \boldsymbol{Y} X_p \\ &= \frac{\sum_{i=1}^{p} d_i^2 B_{i_{(OLS)}}^2}{\sum_{j=1}^{p} d_j^3 B_{j_{(OLS)}}^2} d_1 \cdot B_{1_{(OLS)}} X_1 + \cdots + \frac{\sum_{i=1}^{p} d_i^2 B_{i_{(OLS)}}^2}{\sum_{j=1}^{p} d_j^3 B_{j_{(OLS)}}^2} d_p \cdot B_{p_{(OLS)}} X_p.\end{aligned}$$

That is, the PLS prediction consists of each OLS prediction component scaled by the fraction $\frac{\sum_{i=1}^{p} d_i^2 B_{i_{(OLS)}}^2}{\sum_{j=1}^{p} d_j^3 B_{j_{(OLS)}}^2} d_k$ for each $k = 1, \ldots, p$.

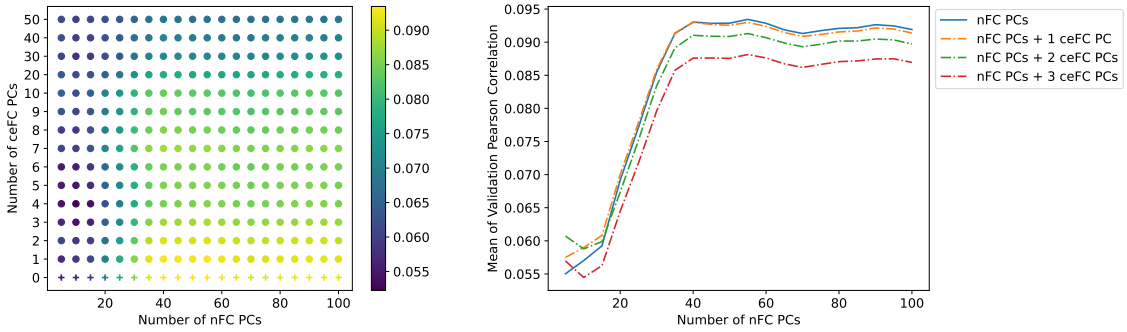
B.2 Linear Regression Validation Results



(a) Validation MSE with all combinations of nFC PCs and ceFC PCs

(b) Validation MSE with selected combinations of nFC PCs and ceFC PCs

Figure B.1: Sum of validation MSEs over all phenotypic traits with different numbers of nFC PCs and ceFC PCs.

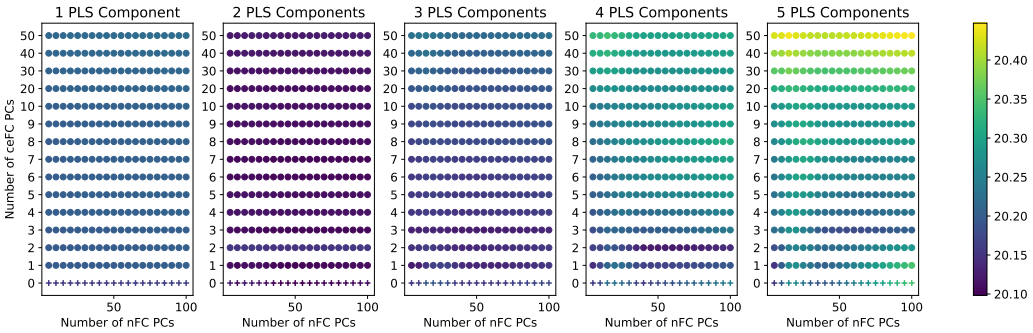


(a) Validation Pearson correlation with all combinations of nFC PCs and ceFC PCs

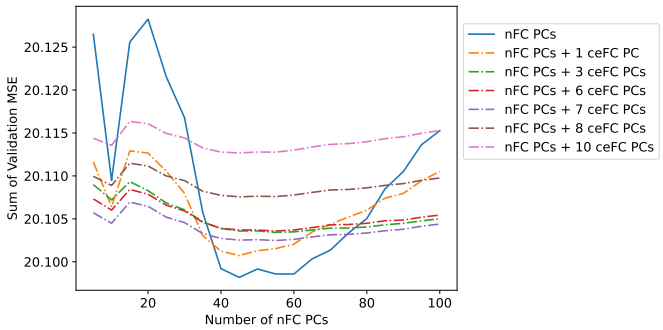
(b) Validation Pearson correlation with selected combinations of nFC PCs and ceFC PCs

Figure B.2: Average validation Pearson correlations over all phenotypic traits with different numbers of nFC PCs and ceFC PCs.

B.3 PLS Regression Validation Results

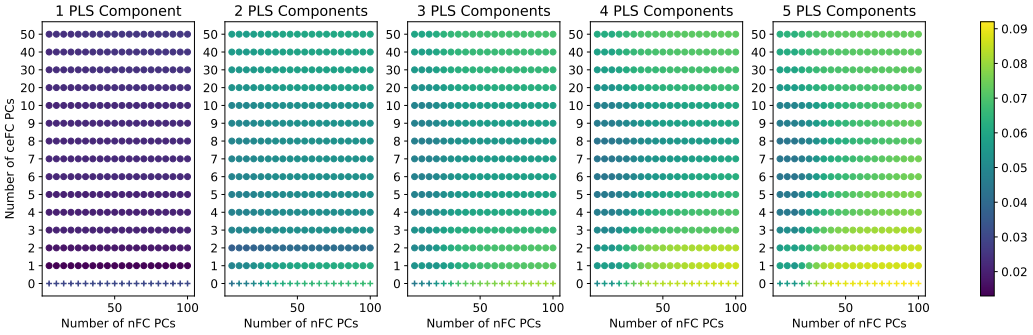


(a) Validation MSE with all combinations of nFC PCs and ceFC PCs under varying number of PLS components.

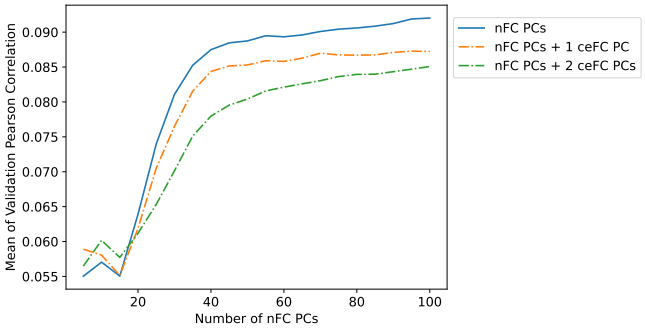


(b) Validation MSE with selected combinations of nFC PC and ceFC PC under 2 PLS components

Figure B.3: Sum of validation MSEs over all phenotypic traits with different numbers of nFC PCs, ceFC PCs, and PLS components.



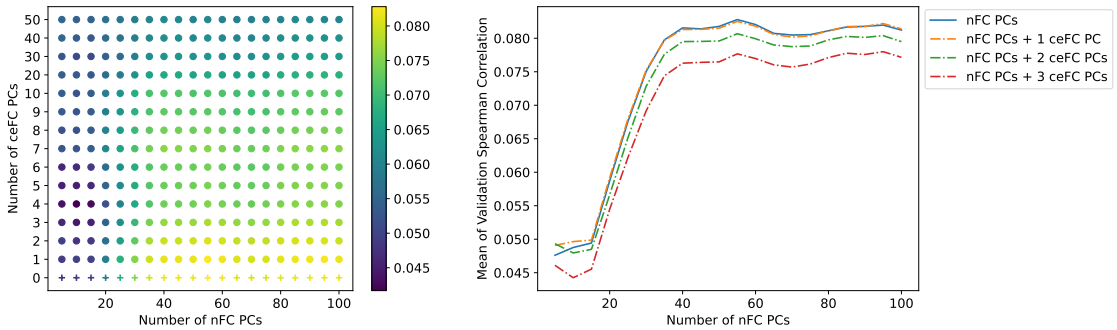
(a) Validation Pearson correlation with all combinations of nFC PCs and ceFC PCs under varying number of PLS components.



(b) Validation Pearson correlation with selected combinations of nFC PCs and ceFC PCs under 5 PLS components

Figure B.4: Average validation Pearson correlations over all phenotypic traits with different numbers of nFC PCs and ceFC PCs.

B.4 Spearman Correlation Results for Linear Regression



(a) Validation Spearman correlation with all combinations of nFC PCs and ceFC PCs

(b) Validation Spearman correlation with selected combinations of nFC PCs and ceFC PCs

Figure B.5: Sum of validation Spearman correlations over all phenotypic traits with different numbers of nFC PCs and ceFC PCs.

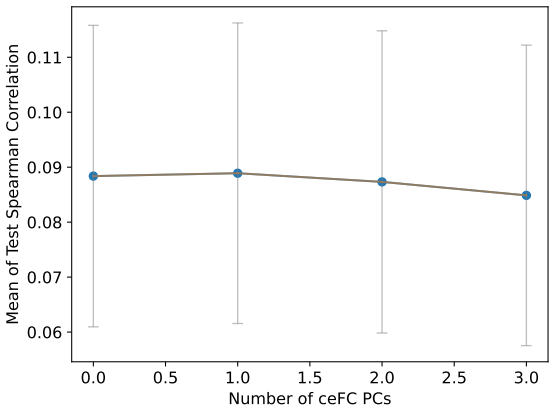


Figure B.6: Mean test Spearman correlation using 55 nFC PCs and varying numbers of ceFC PCs. The gray bars depict standard errors across 100 random splits.

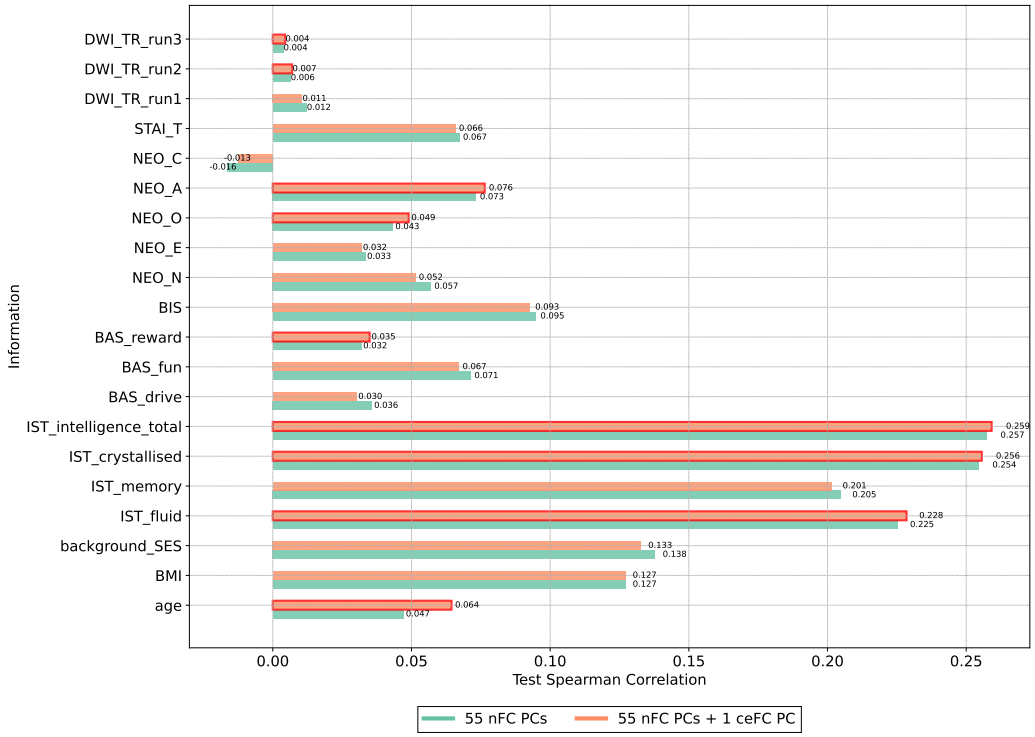
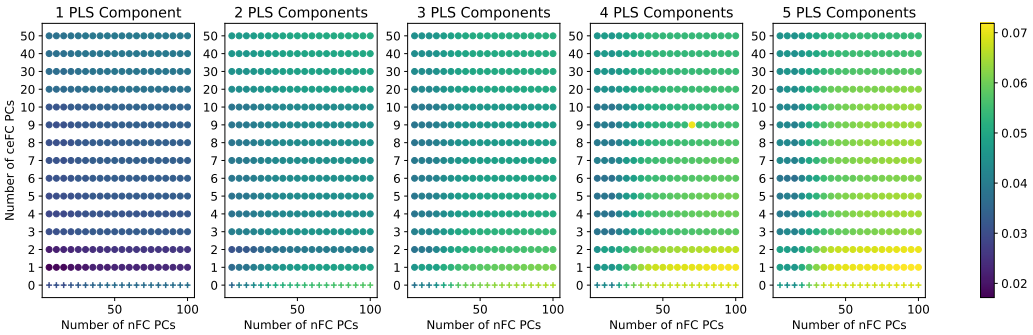
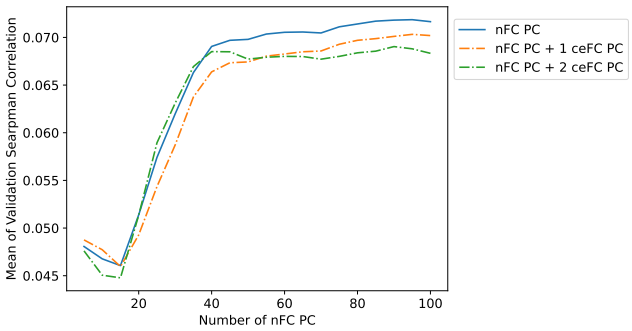


Figure B.7: Test Spearman correlation for each phenotypic trait. The green and orange bars represent results using 55 FC PCs only and 55 FC PCs with 1 ceFC PC, respectively. The highlighted red bar denotes the case when adding the ceFC PC returns the largest test correlation.

B.5 Spearman Correlation Results for PLS Regression



(a) Validation Spearman correlation with all combinations of nFC PCs and ceFC PCs under varying numbers of PLS components.



(b) Validation Spearman correlation with selected combinations of nFC PCs and ceFC PCs under 5 PLS components

Figure B.8: Mean of validation Spearman correlations over all phenotypic traits with different numbers of nFC PCs and ceFC PCs.

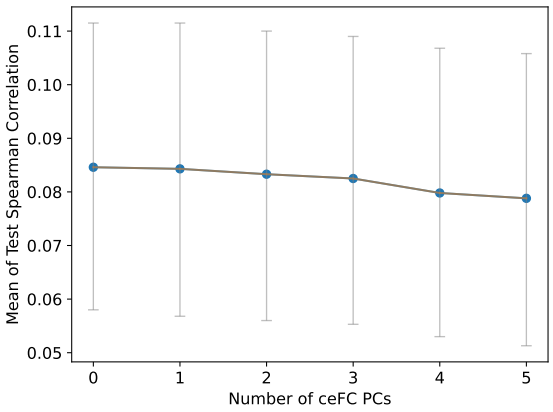


Figure B.9: Mean test Spearman correlation using 100 nFC PCs and varying numbers of ceFC PCs under 5 PLS components. The gray bars depict standard errors across 100 random splits.

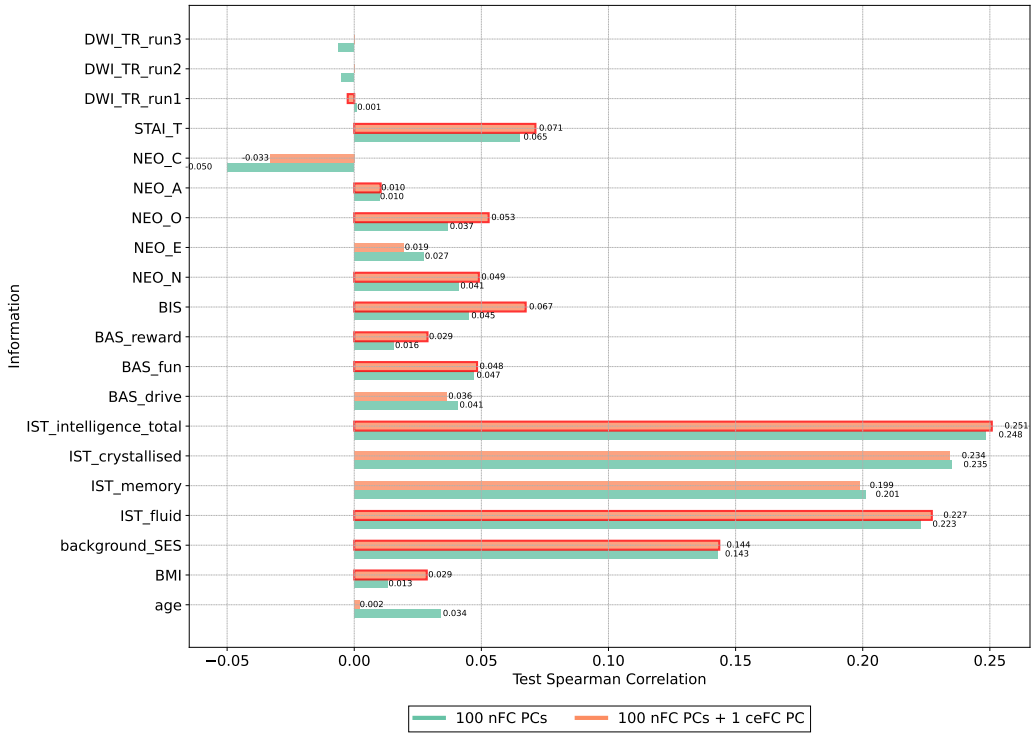


Figure B.10: Test Spearman correlation for each phenotypic trait with 5 PLS components. The green and orange bars represent results using 100 nFC PCs only and 100 nFC PCs with 1 ceFC PC, respectively. The highlighted red bar denotes the case when adding the ceFC PCd returns the largest test correlation.

APPENDIX C

Appendix of Chapter III

C.1 Proof of Proposition 4.1

Proof. Given $\pi\{\mathbf{y}(\mathbf{s},t) \mid \mathbf{X}(t), \widetilde{\boldsymbol{\lambda}}(\mathbf{s},t)\} = \pi\{\mathbf{y}(\mathbf{s},t) \mid \mathbf{X}(t), \bar{\boldsymbol{\lambda}}(\mathbf{s},t)\}$ for all $\mathbf{s}$ and t, without loss of generality, we fix the time at $\check{t}$ and obtain

$$\mathbb{E}[\mathbf{y}(\mathbf{s},\check{t}) \mid \mathbf{X}(\check{t}), \widetilde{\boldsymbol{\lambda}}(\mathbf{s},\check{t})] = \mathbb{E}[\mathbf{y}(\mathbf{s},\check{t}) \mid \mathbf{X}(\check{t}), \bar{\boldsymbol{\lambda}}(\mathbf{s},\check{t})].$$

Equivalently,

$$\mathbf{X}(\check{t})\widetilde{\boldsymbol{\beta}}(\mathbf{s},\check{t}) + \widetilde{\boldsymbol{\alpha}}(\mathbf{s},\check{t}) = \mathbf{X}(\check{t})\bar{\boldsymbol{\beta}}(\mathbf{s},\check{t}) + \bar{\boldsymbol{\alpha}}(\mathbf{s},\check{t}).$$

Then

$$\begin{aligned}
\mathbf{X}(\check{t})\ [\widetilde{\boldsymbol{\beta}}(\mathbf{s},\check{t}) - \bar{\boldsymbol{\beta}}(\mathbf{s},\check{t})\ ] &= \bar{\boldsymbol{\alpha}}(\mathbf{s},\check{t}) - \widetilde{\boldsymbol{\alpha}}(\mathbf{s},\check{t}),\\
\widetilde{\boldsymbol{\beta}}(\mathbf{s},\check{t}) - \bar{\boldsymbol{\beta}}(\mathbf{s},\check{t}) &= (\mathbf{X}(\check{t})^\top\mathbf{X}(\check{t}))^{-1}\mathbf{X}(\check{t})^\top\bar{\boldsymbol{\alpha}}(\mathbf{s},\check{t}) - (\mathbf{X}(\check{t})^\top\mathbf{X}(\check{t}))^{-1}\mathbf{X}(\check{t})^\top\widetilde{\boldsymbol{\alpha}}(\mathbf{s},\check{t})\\
&= \mathbf{0} - \mathbf{0} = \mathbf{0},
\end{aligned}$$

by $\mathbf{X}(\check{t})^T\boldsymbol{\alpha}(\mathbf{s},\check{t}) = \mathbf{0}$ for all $\mathbf{s} \in \mathcal{S}$.

Furthermore,

$$\widetilde{\boldsymbol{\alpha}}(\mathbf{s},\check{t}) - \bar{\boldsymbol{\alpha}}(\mathbf{s},\check{t}) = \mathbf{X}(\check{t})\ [\bar{\boldsymbol{\beta}}(\mathbf{s},\check{t}) - \widetilde{\boldsymbol{\beta}}(\mathbf{s},\check{t})\ ] = \mathbf{0}.$$

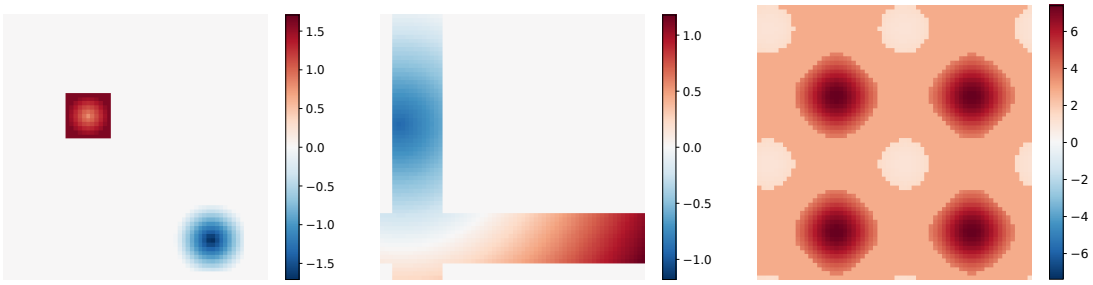
Finally,

$$\tilde{\sigma}^2(\mathbf{s},\check{t}) = \mathrm{Var}(y_1(\mathbf{s},\check{t}) \mid \mathbf{X}(\check{t}), \widetilde{\boldsymbol{\lambda}}(\mathbf{s},\check{t})) = \mathrm{Var}(y_1(\mathbf{s},\check{t}) \mid \mathbf{X}(\check{t}), \bar{\boldsymbol{\lambda}}(\mathbf{s},\check{t})) = \bar{\sigma}^2(\mathbf{s},\check{t}).$$

Therefore, we have $\widetilde{\boldsymbol{\lambda}}(\mathbf{s},\check{t}) = \bar{\boldsymbol{\lambda}}(\mathbf{s},\check{t})$ for any $\check{t}$. Combining the result for any time $t \in [0,1]$, we obtain $\widetilde{\boldsymbol{\lambda}}(\mathbf{s},t) = \bar{\boldsymbol{\lambda}}(\mathbf{s},t)$ for all $\mathbf{s}$ and t.

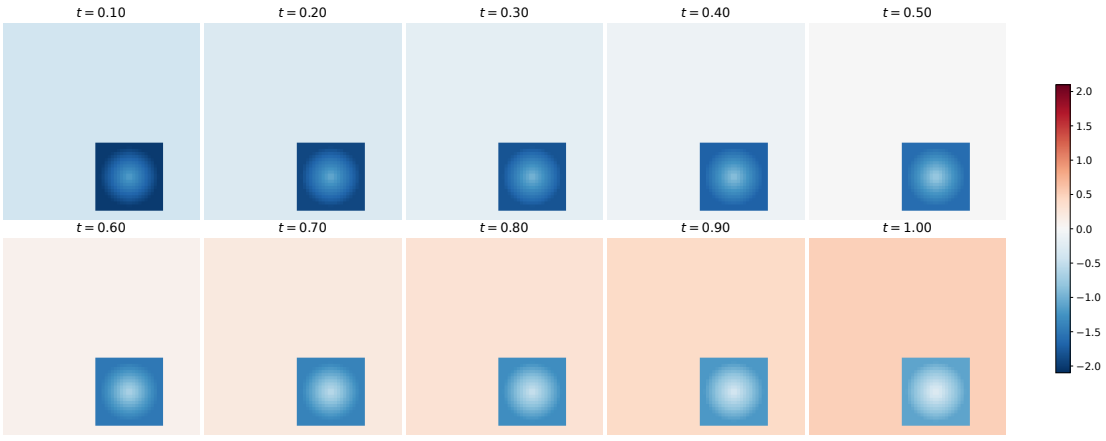
□

C.2 Examples of Spatio-temporal Design in Numerical Studies

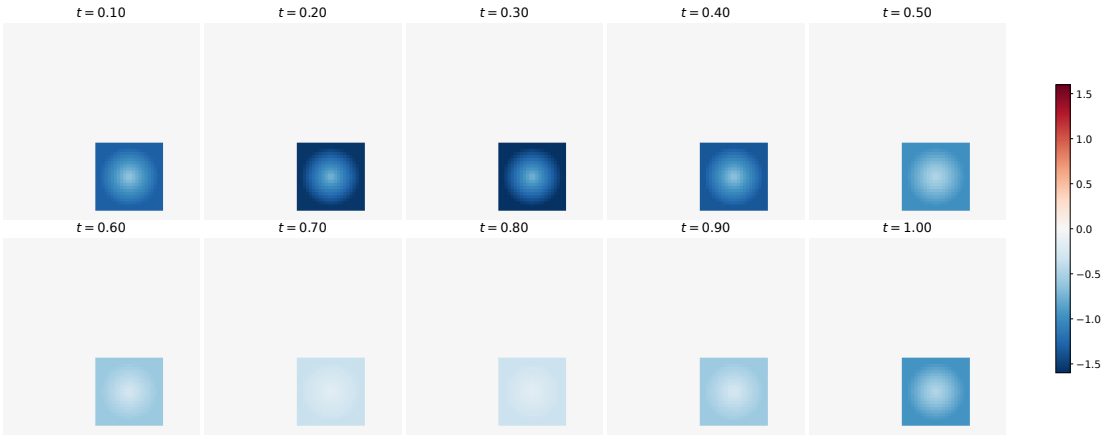


(a) Main effect (b) Individual effect (c) Log variance of error

Figure C.1: A spatial design example for (a) a main effect $\beta_1(\boldsymbol{s}_v)$, (b) an individual effect $\alpha_1(\boldsymbol{s}_v)$, and (c) log variance of the error term $\log[\sigma^2(\boldsymbol{s}_v)]$.

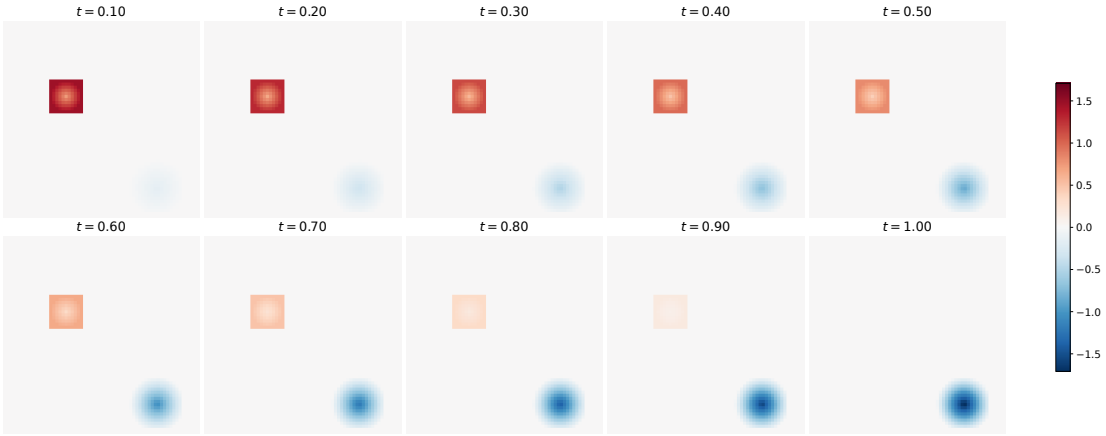


(a) Additive spatio-temporal design

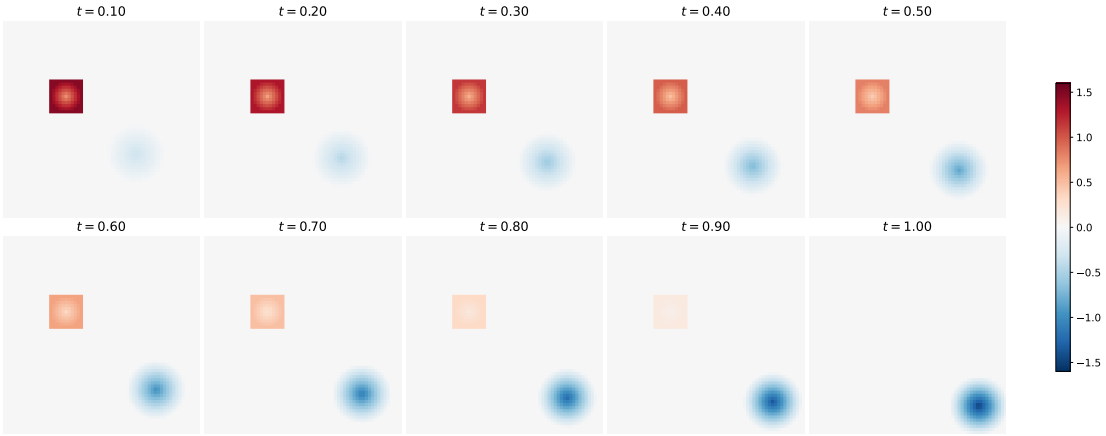


(b) Multiplicative spatio-temporal design

Figure C.2: Examples of spatio-temporal designs, continued in the next figure.



(c) Concatenative spatio-temporal design



(d) Concatenative spatio-temporal design with a moving activation region

Figure C.2: Examples of spatio-temporal designs for (a) additive, (b) multiplicative, (c) concatenative, and (d) concatenative with a moving activation region.

BIBLIOGRAPHY

Achard, S. and Bullmore, E. (2007). Efficiency and Cost of Economical Brain Functional Networks. *PLOS Computational Biology* **3** e17.

Ackerman, S. (1992). *Discovering the Brain*. National Academies Press (US).

Aertsen, A., Gerstein, G., Habib, M. and Palm, G. (1989). Dynamics of neuronal firing correlation: modulation of" effective connectivity". *Journal of neurophysiology* **61** 900–917.

Amthauer, R., Brocke, B., Liepmann, D. and Beauducel, A. (2001). *IST 2000 R: Intelligenz-Struktur-Test 2000 R*. Hogrefe, Verlag für Psychologie Bern, Switzerland.

Asadi, N., Olson, I. R. and Obradovic, Z. (2023). A transformer model for learning spatiotemporal contextual representation in fmri data. *Network Neuroscience* **7** 22–47.

Asai, M., Gupta, R. and McAleer, M. (2020). Forecasting volatility and covolatility of crude oil and gold futures: Effects of leverage, jumps, spillovers, and geopolitical risks. *International Journal of Forecasting* **36** 933–948.

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U. and Norman, K. A. (2017). Discovering event structure in continuous narrative perception and memory. *Neuron* **95** 709–721.e5.

Barut, E., Fan, J. and Verhasselt, A. (2016). Conditional sure independence screening. *Journal of the American Statistical Association* **111** 1266–1277.

Beckmann, C. F., DeLuca, M., Devlin, J. T. and Smith, S. M. (2005). Investigations into resting-state connectivity using independent component analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360** 1001–1013.

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** 289–300.

Benjamini, Y. and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics* **29** 1165 – 1188.

Bertolero, M. A., Yeo, B. T., Bassett, D. S. and D'Esposito, M. (2018). A mechanistic model of connector hubs, modularity and cognition. *Nature human behaviour* **2** 765–777.

Bickel, P. J. and Levina, E. (2008a). Covariance regularization by thresholding. *The Annals of Statistics* **36** 2577–2604.

Bickel, P. J. and Levina, E. (2008b). Regularized estimation of large covariance matrices. *The Annals of Statistics* **36** 199–227.

Bien, J., Bunea, F. and Xiao, L. (2016). Convex banding of the covariance matrix. *Journal of the American Statistical Association* **111** 834–845. PMID: 28042189.

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

Bryce, N. V., Flournoy, J. C., Guassi Moreira, J. F., Rosen, M. L., Sambook, K. A., Mair, P. and McLaughlin, K. A. (2021). Brain parcellation selection: An overlooked decision point with meaningful effects on individual differences in resting-state functional connectivity. *NeuroImage* **243** 118487.

Buchanan, C. R., Bastin, M. E., Ritchie, S. J., Liewald, D. C., Madole, J. W., Tucker-Drob, E. M., Deary, I. J. and Cox, S. R. (2020). The effect of network thresholding and weighting on structural brain networks in the uk biobank. *NeuroImage* **211** 116443.

Buckner, R. L., Bandettini, P. A., O'Craven, K. M., Savoy, R. L., Petersen, S. E., Raichle, M. E. and Rosen, B. R. (1996). Detection of cortical activation during averaged single trials of a cognitive task using functional magnetic resonance imaging. *Proceedings of the National Academy of Sciences* **93** 14878–14883.

Bulik-Sullivan, B. K., Finucane, H. K., Anttila, V., Gusev, A., Day, F. R., Loh, P.-R., Duncan, L. E., Perry, J. R. B., Patterson, N. J., Robinson, E. B., Daly, M. J., Price, A. L. and Neale, B. M. (2015). An atlas of genetic correlations across human diseases and traits. *Nature genetics* **47** 1236 – 1241.

Bullmore, E. and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature reviews neuroscience* **10** 186–198.

Bullmore, E. and Sporns, O. (2012). The economy of brain network organization. *Nature reviews Neuroscience* **13** 336–349.

Bzdok, D., Varoquaux, G. and Steyerberg, E. W. (2021). Prediction, not association, paves the road to precision medicine. *JAMA Psychiatry* **78** 127–128.

Cai, T. and Liu, W. (2011). Adaptive thresholding for sparse covariance matrix estimation. *Journal of the American Statistical Association* **106** 672–684.

Cai, T. T., Zhang, C.-H. and Zhou, H. H. (2010). Optimal rates of convergence for covariance matrix estimation. *The Annals of Statistics* **38** 2118 – 2144.

Callejas, A., Shulman, G. L. and Corbetta, M. (2014). Dorsal and ventral attention systems underlie social and symbolic cueing. *J. Cognitive Neuroscience* **26** 63–80.

Carver, C. S. and White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the bis/bas scales. *Journal of personality and social psychology* **67** 319.

Chen, J., Leong, Y. C., Honey, C. J., Yong, C. H., Norman, K. A. and Hasson, U. (2017). Shared memories reveal shared structure in neural activity across individuals. *Nature neuroscience* **20** 115–125.

Chen, J., Tam, A., Kebets, V., Orban, C., Ooi, L., Asplund, C., Marek, S., Dosenbach, N., Eickhoff, S., Bzdok, D., Holmes, A. and Yeo, B. (2022). Shared and unique brain network features predict cognitive, personality, and mental health scores in the abcd study. *Nature Communications* **13**.

Chen, X., Zhang, H., Gao, Y., Wee, C.-Y., Li, G., Shen, D. and the Alzheimer's Disease Neuroimaging Initiative (2016). High-order resting-state functional connectivity network for MCI classification. *Human Brain Mapping* **37** 3282–3296.

Chen, Z. J., He, Y., Rosa-Neto, P., Germann, J. and Evans, A. C. (2008). Revealing modular architecture of human brain structural networks by using cortical thickness from mri. *Cerebral cortex* **18** 2374–2381.

Cheng, W., Palaniyappan, L., Li, M., Kendrick, K. M., Zhang, J., Luo, Q., Liu, Z., Yu, R., Deng, W., Wang, Q. et al. (2015). Voxel-based, brain-wide association study of aberrant functional connectivity in schizophrenia implicates thalamocortical circuitry. *npj Schizophrenia* **1** 1–8.

Chenji, S., Jha, S., Lee, D., Brown, M., Seres, P., Mah, D. and Kalra, S. (2016). Investigating Default Mode and Sensorimotor Network Connectivity in Amyotrophic Lateral Sclerosis. *PLOS ONE* **11** 1–14.

Chumin, E. J., Faskowitz, J., Esfahlani, F. Z., Jo, Y., Merritt, H., Tanner, J., Cutts, S. A., Pope, M., Betzel, R. and Sporns, O. (2022). Cortico-subcortical interactions in overlapping communities of edge functional connectivity. *NeuroImage* **250** 118971.

Corbetta, M. and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience* **3** 201–215.

Cressie, N. and Wikle, C. K. (2011). *Statistics for spatio-temporal data*. John Wiley & Sons.

Dai, P., Zhou, Y., Shi, Y., Lu, D., Chen, Z., Zou, B., Liu, K., Liao, S. and meta MDD Consortium, T. R. (2024). Classification of mdd using a transformer classifier with large-scale multisite resting-state fmri data. *Human Brain Mapping* **45** e26542.

Dai, W., Qiu, E., Lin, X., Zhang, S., Zhang, M., Han, X., Jia, Z., Su, H., Bian, X., Zang, X., Li, M., Zhang, Q., Ran, Y., Gong, Z., Wang, X., Wang, R., Tian, L., Dong, Z. and Yu, S. (2023). Abnormal Thalamo–Cortical Interactions in Overlapping Communities of Migraine: An Edge Functional Connectivity Study. *Annals of Neurology* **94** 1168–1181.

de Reus, M. A. and van den Heuvel, M. P. (2013). Estimating false positives and negatives in brain networks. *NeuroImage* **70** 402–409.

De Vico Fallani, F., Latora, V. and Chavez, M. (2017). A Topological Criterion for Filtering Information in Complex Brain Networks. *PLOS Computational Biology* **13** e1005305.

Deng, X., Zhang, J., Liu, R. and Liu, K. (2022). Classifying asd based on time-series fmri using spatial–temporal transformer. *Computers in Biology and Medicine* **151** 106320.

Devlin, J., Chang, M.-W., Lee, K. and Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. In *North American Chapter of the Association for Computational Linguistics*.

Dosenbach, N. U., Fair, D. A., Cohen, A. L., Schlaggar, B. L. and Petersen, S. E. (2008). A dual-networks architecture of top-down control. *Trends in cognitive sciences* **12** 99–105.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J. and Houlsby, N. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. In *International Conference on Learning Representations*.

Duncan, L., Yilmaz, Z., Gaspar, H., Walters, R., Goldstein, J., Anttila, V., Bulik-Sullivan, B., Ripke, S., of the Psychiatric Genomics Consortium, E. D. W. G., Thornton, L. et al. (2017). Significant locus and metabolic genetic correlations revealed in genome-wide association study of anorexia nervosa. *American journal of psychiatry* **174** 850–858.

Esfahlani, F. Z., Jo, Y., Faskowitz, J., Byrge, L., Kennedy, D. P., Sporns, O. and Betzel, R. F. (2020). High-amplitude cofluctuations in cortical activity drive functional connectivity. *Proceedings of the National Academy of Sciences* **117** 28393–28401.

Fan, J. and Liu, H. (2013). Statistical analysis of big data on pharmacogenomics. *Advanced drug delivery reviews* **65** 987–1000.

Faskowitz, J., Betzel, R. F. and Sporns, O. (2022). Edges in brain networks: Contributions to models of structure and function. *Network Neuroscience* **6** 1–28.

Faskowitz, J., Esfahlani, F. Z., Jo, Y., Sporns, O. and Betzel, R. F. (2020). Edge-centric functional network representations of human cerebral cortex reveal overlapping system-level architecture. *Nature neuroscience* **23** 1644—1654.

Finn, E. S. and Bandettini, P. A. (2021). Movie-watching outperforms rest for functional connectivity-based prediction of behavior. *NeuroImage* **235** 117963.

Finn, E. S. and Constable, R. T. (2016). Individual variation in functional brain connectivity: implications for personalized approaches to psychiatric disease. *Dialogues in Clinical Neuroscience* **18** 277–287.

Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., Papademetris, X. and Constable, R. T. (2015). Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nature neuroscience* **18** 1664–1671.

Fox, M. D., Snyder, A. Z., Vincent, J. L. and Raichle, M. E. (2007). Intrinsic fluctuations within cortical systems account for intertrial variability in human behavior. *Neuron* **56** 171–184.

Franken, I. H., Muris, P. and Rassin, E. (2005). Psychometric properties of the dutch bis/bas scales. *Journal of Psychopathology and Behavioral Assessment* **27** 25–30.

Friston, K. J. (1994). Functional and effective connectivity in neuroimaging: a synthesis. *Human brain mapping* **2** 56–78.

Friston, K. J. (2003). Statistical parametric mapping. In *Neuroscience databases: a practical guide*. Springer, 237–250.

Friston, K. J., Frith, C. D., Liddle, P. F. and Frackowiak, R. S. J. (1993). Functional connectivity: The principal-component analysis of large (pet) data sets. *Journal of Cerebral Blood Flow & Metabolism* **13** 5–14. PMID: 8417010.

Fulwiler, C. E., King, J. A. and Zhang, N. (2012). Amygdala–orbitofrontal resting-state functional connectivity is associated with trait anger. *Neuroreport* **23** 606–610.

Gallos, L. K., Makse, H. A. and Sigman, M. (2012). A small world of weak ties provides optimal global integration of self-similar modules in functional brain networks. *Proceedings of the National Academy of Sciences* **109** 2825–2830.

Gao, J., Ge, B., Qiang, N. and Zhao, S. (2026). 3d masked autoencoder with spatiotemporal transformer for modeling of 4d fmri data. *Medical Image Analysis* **107** 103861.

Garrison, K. A., Scheinost, D., Finn, E. S., Shen, X. and Constable, R. T. (2015). The (in)stability of functional brain network measures across thresholds. *NeuroImage* **118** 651–661.

Gatica, M., Cofré, R., Mediano, P. A., Rosas, F. E., Orio, P., Diez, I., Swinnen, S. P. and Cortes, J. M. (2021). High-order interdependencies in the aging brain. *Brain Connectivity* **11** 734–744. PMID: 33858199.

Geerligs, L., Renken, R. J., Saliasi, E., Maurits, N. M. and Lorist, M. M. (2015). A brain-wide study of age-related changes in functional connectivity. *Cerebral cortex* **25** 1987–1999.

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C. F., Jenkinson, M. et al. (2016). A multi-modal parcellation of human cerebral cortex. *Nature* **536** 171–178.

Glickfeld, L. L. and Olsen, S. R. (2017). Higher-order areas of the mouse visual cortex. *Annual review of vision science* **3** 251–273.

Gordon, E. M., Laumann, T. O., Adeyemo, B., Huckins, J. F., Kelley, W. M. and Petersen, S. E. (2016). Generation and evaluation of a cortical area parcellation from resting-state correlations. *Cerebral cortex* **26** 288–303.

Gordon, E. M., Laumann, T. O., Gilmore, A. W., Newbold, D. J., Greene, D. J., Berg, J. J., Ortega, M., Hoyt-Drazen, C., Gratton, C., Sun, H., Hampton, J. M., Coalson, R. S., Nguyen, A. L., McDermott, K. B., Shimony, J. S., Snyder, A. Z., Schlaggar, B. L., Petersen, S. E., Nelson, S. M. and Dosenbach, N. U. (2017). Precision functional mapping of individual human brains. *Neuron* **95** 791–807.e7.

Gratton, C., Nelson, S. M. and Gordon, E. M. (2022). Brain-behavior correlations: Two paths toward reliability. *Neuron* **110** 1446–1449.

Greicius, M. D., Krasnow, B., Reiss, A. L. and Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proceedings of the national academy of sciences* **100** 253–258.

Hampson, M., Driesen, N. R., Skudlarski, P., Gore, J. C. and Constable, R. T. (2006). Brain connectivity related to working memory performance. *Journal of Neuroscience* **26** 13338–13343.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G. and Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science* **303** 1634–1640.

Hasson, U., Yang, E., Vallines, I., Heeger, D. J. and Rubin, N. (2008). A hierarchy of temporal receptive windows in human cortex. *Journal of neuroscience* **28** 2539–2550.

Helmer, M., Warrington, S., Mohammadi-Nejad, A.-R., Ji, J. L., Howell, A., Rosand, B., Anticevic, A., Sotiropoulos, S. N. and Murray, J. D. (2024). On stability of canonical correlation analysis and partial least squares with application to brain-behavior associations. *Communications Biology* .

Honey, C. J., Kötter, R., Breakspear, M. and Sporns, O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Sciences* **104** 10240–10245.

Honey, C. J., Thesen, T., Donner, T. H., Silbert, L. J., Carlson, C. E., Devinsky, O., Doyle, W. K., Rubin, N., Heeger, D. J. and Hasson, U. (2012). Slow cortical dynamics and the accumulation of information over long timescales. *Neuron* **76** 423–434.

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

Hove, M. and Martinez, S. A. (2024). *Biological Psychology [Revised Edition]*. ROTEL (Remixing Open Textbooks with an Equity Lens) Project.

Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E. and Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* **532** 453–458.

Isserlis, L. (1918). On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika* **12** 134–139.

Jo, Y., Faskowitz, J., Esfahlani, F. Z., Sporns, O. and Betzel, R. F. (2021a). Subject identification using edge-centric functional connectivity. *NeuroImage* **238** 118204.

Jo, Y., Zamani Esfahlani, F., Faskowitz, J., Chumin, E. J., Sporns, O. and Betzel, R. F. (2021b). The diversity and multiplexity of edge communities within and between brain systems. *Cell reports* **37** 110032.

Johnstone, I. M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *The Annals of Statistics* **29** 295–327.

Johnstone, I. M. and Lu, A. Y. (2009). On consistency and sparsity for principal components analysis in high dimensions. *Journal of the American Statistical Association* **104** 682–693. PMID: 20617121.

Kaiser, R. H., Andrews-Hanna, J. R., Wager, T. D. and Pizzagalli, D. A. (2015). Large-Scale Network Dysfunction in Major Depressive Disorder: A Meta-analysis of Resting-State Functional Connectivity. *JAMA Psychiatry* **72** 603–611.

Kang, K., Seidlitz, J., Bethlehem, R. A., Xiong, J., Jones, M. T., Mehta, K., Keller, A. S., Tao, R., Randolph, A., Larsen, B. et al. (2024). Study design features increase replicability in brain-wide association studies. *Nature* **636** 719–727.

Karoui, N. E. (2008). Operator norm consistent estimation of large-dimensional sparse covariance matrices. *The Annals of Statistics* **36** 2717 – 2756.

Ke, J., Song, H., Bai, Z., Rosenberg, M. D. and Leong, Y. C. (2025). Dynamic brain connectivity predicts emotional arousal during naturalistic movie-watching. *PLOS Computational Biology* **21** 1–26.

Khambhati, A. N., Shafi, A., Rao, V. R. and Chang, E. F. (2021). Long-term brain network reorganization predicts responsive neurostimulation outcomes for focal epilepsy. *Science Translational Medicine* **13** eabf6588.

Kim, P. Y., Kwon, J., Joo, S., Bae, S., Lee, D., Jung, Y., Yoo, S., Cha, J. and Moon, T. (2023). Swift: swin 4d fmri transformer. In *Proceedings of the 37th International Conference on Neural Information Processing Systems*. NIPS '23, Curran Associates Inc., Red Hook, NY, USA.

Kinsey, S., Kazimierczak, K., Camazón, P., Chen, J., Adali, T., Kochunov, P., Adhikari, B., Ford, J., Erp, T., Dhamala, M., Calhoun, V. and Iraji, A. (2024). Networks extracted from nonlinear fmri connectivity exhibit unique spatial variation and enhanced sensitivity to differences between individuals with schizophrenia and controls. *Nature Mental Health* **2** 1464–1475.

Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A. and Poeppel, D. (2017). Neuroscience needs behavior: correcting a reductionist bias. *Neuron* **93** 480–490.

Krishnan, A., Williams, L. J., McIntosh, A. R. and Abdi, H. (2011). Partial least squares (pls) methods for neuroimaging: a tutorial and review. *Neuroimage* **56** 455–475.

Kwong, K. K., Belliveau, J. W., Chesler, D. A., Goldberg, I. E., Weisskoff, R. M., Poncelet, B. P., Kennedy, D. N., Hoppel, B. E., Cohen, M. S. and Turner, R. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences* **89** 5675–5679.

LaBar, K. S., Gitelman, D. R., Parrish, T. B., Kim, Y.-H., Nobre, A. C., Mesulam, M. et al. (2001). Hunger selectively modulates corticolimbic activation to food stimuli in humans. *Behavioral neuroscience* **115** 493.

Lahaye, P.-J., Poline, J.-B., Flandin, G., Dodel, S. and Garnero, L. (2003). Functional connectivity: studying nonlinear, delayed interactions between bold signals. *NeuroImage* **20** 962–974.

Lahnakoski, J. M., Glerean, E., Salmi, J., Jääskeläinen, I. P., Sams, M., Hari, R. and Nummenmaa, L. (2012). Naturalistic fmri mapping reveals superior temporal sulcus as the hub for the distributed brain network for social perception. *Frontiers in Human Neuroscience* **Volume 6 - 2012**.

Lam, C. and Fan, J. (2009). Sparsistency and rates of convergence in large covariance matrix estimation. *The Annals of Statistics* **37** 4254 – 4278.

Lerner, Y., Honey, C. J., Silbert, L. J. and Hasson, U. (2011). Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *Journal of Neuroscience* **31** 2906–2915.

Lettieri, G., Handjaras, G., Setti, F., Cappello, E. M., Bruno, V., Diano, M., Leo, A., Ricciardi, E., Pietrini, P. and Cecchetti, L. (2022). Default and control network connectivity dynamics track the stream of affect at multiple timescales. *Social Cognitive and Affective Neuroscience* **17** 461–469.

Li, J., Huang, C. and and, Z. H. (2017). A functional varying-coefficient single-index model for functional response data. *Journal of the American Statistical Association* **112** 1169–1181. PMID: 29200540.

Li, X., Wang, L. and and, H. J. W. (2021). Sparse learning and structure identification for ultrahigh-dimensional image-on-scalar regression. *Journal of the American Statistical Association* **116** 1994–2008.

Liang, J., Cao, J., Sun, G., Zhang, K., Van Gool, L. and Timofte, R. (2021). Swinir: Image restoration using swin transformer. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*.

Liebenthal, E. and Möttönen, R. (2018). An interactive model of auditory-motor speech perception. *Brain and Language* **187** 33–40.

Liepmann, D., Beauducel, A., Brocke, B., Amthauer, R. and Vorst, H. (2010). Intelligentie structuur test. *Hogrefe Uitgevers BV Dutch Translation by Vorst HCM* .

Liu, S., Abdellaoui, A., Verweij, K. J. and van Wingen, G. A. (2023). Replicable brain–phenotype associations require large-scale neuroimaging data. *Nature Human Behaviour* **7** 1344–1356.

Liu, Y., Fang, E. X. and Lu, J. (2022a). Lagrangian inference for ranking problems. *Operations Research* .

Liu, Z.-Q., Vazquez-Rodriguez, B., Spreng, R. N., Bernhardt, B. C., Betzel, R. F. and Misic, B. (2022b). Time-resolved structure-function coupling in brain networks. *Communications Biology* **5** 532.

Luo, T., Ma, Z., Wang, Z., John Xu, Z. and Zhang, Y. (2022). An upper limit of decaying rate with respect to frequency in linear frequency principle model. In *Proceedings of Mathematical and Scientific Machine Learning* (B. Dong, Q. Li, L. Wang and Z.-Q. J. Xu, eds.), vol. 190 of *Proceedings of Machine Learning Research*. PMLR.

Luppi, A. I., Mediano, P. A., Rosas, F. E., Holland, N., Fryer, T. D., O'Brien, J. T., Rowe, J. B., Menon, D. K., Bor, D. and Stamatakis, E. A. (2022). A synergistic core for human brain evolution and cognition. *Nature Neuroscience* **25** 771–782.

Magnus, J. R. and Neudecker, H. (1979). The Commutation Matrix: Some Properties and Applications. *The Annals of Statistics* **7** 381 – 394.

Mantwill, M., Gell, M., Krohn, S. and Finke, C. (2022). Brain connectivity fingerprinting and behavioural prediction rest on distinct functional systems of the human connectome. *Communications biology* **5** 261.

Marčenko, V. A. and Pastur, L. A. (1967). Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik* **1** 457.

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** 654–660.

Markon, K. E. (2010). Modeling psychopathology structure: a symptom-level analysis of Axis I and II disorders. *Psychological Medicine* **40** 273–288.

Martynenko, I., Koldewyn, K. and Downing, P. E. (2025). Social brain responses to natural scene images of social interactions. *Social Cognitive and Affective Neuroscience* **20** nsaf057.

Matthews, B. W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta (BBA)-Protein Structure* **405** 442–451.

Matusz, P. J., Dikker, S., Huth, A. G. and Perrodin, C. (2019). Are we ready for real-world neuroscience? *Journal of Cognitive Neuroscience* **31** 327–338.

McIntosh, A., Bookstein, F., Haxby, J. V. and Grady, e. C. (1996). Spatial pattern analysis of functional brain images using partial least squares. *Neuroimage* **3** 143–157.

McMahon, E. and Isik, L. (2023). Seeing social interactions. *Trends in Cognitive Sciences* **27** 1165–1179.

Monk, C. S., Peltier, S. J., Wiggins, J. L., Weng, S.-J., Carrasco, M., Risi, S. and Lord, C. (2009). Abnormalities of intrinsic functional connectivity in autism spectrum disorders. *Neuroimage* **47** 764–772.

Motlaghian, S. M., Belger, A., Bustillo, J. R., Ford, J. M., Iraji, A., Lim, K., Mathalon, D. H., Mueller, B. A., O'Leary, D., Pearlson, G., Potkin, S. G., Preda, A., van Erp, T. G. M. and Calhoun, V. D. (2022). Nonlinear functional network connectivity in resting functional magnetic resonance imaging data. *Human Brain Mapping* **43** 4556–4566.

Mueller, S., Wang, D., Fox, M., Yeo, B., Sepulcre, J., Sabuncu, M., Shafee, R., Lu, J. and Liu, H. (2013). Individual variability in functional connectivity architecture of the human brain. *Neuron* **77** 586–595.

Nicolini, C., Forcellini, G., Minati, L. and Bifone, A. (2020). Scale-resolved analysis of brain functional connectivity networks with spectral entropy. *NeuroImage* **211** 116603.

Niu, K., Guo, J., Pan, Y., Gao, X., Peng, X., Li, N. and Li, H. (2020). Multichannel deep attention neural networks for the classification of autism spectrum disorder using neuroimaging and personal characteristic data. *Complexity* **2020** 1357853.

Nummenmaa, L., Glerean, E., Viinikainen, M., Jääskeläinen, I. P., Hari, R. and Sams, M. (2012). Emotions promote social interaction by synchronizing brain activity across individuals. *Proceedings of the National Academy of Sciences* **109** 9599–9604.

Ogawa, S., Lee, T.-M., Kay, A. R. and Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences* **87** 9868–9872.

Ooi, L. Q. R., Chen, J., Zhang, S., Kong, R., Tam, A., Li, J., Dhamala, E., Zhou, J. H., Holmes, A. J. and Yeo, B. T. (2022a). Comparison of individualized behavioral predictions across anatomical, diffusion and functional connectivity mri. *NeuroImage* **263** 119636.

Ooi, L. Q. R., Chen, J., Zhang, S., Kong, R., Tam, A., Li, J., Dhamala, E., Zhou, J. H., Holmes, A. J. and Yeo, B. T. T. (2022b). Comparison of individualized behavioral predictions across anatomical, diffusion and functional connectivity mri. *NeuroImage* **263** 119636.

Palla, G., Derényi, I., Farkas, I. and Vicsek, T. (2005). Uncovering the overlapping community structure of complex networks in nature and society. *Nature* **435** 814–818.

Pan, N., Wang, S., Lan, H., Zhang, X., Qin, K., Kemp, G., Suo, X. and Gong, Q. (2024). Multivariate patterns of brain functional connectome associated with covid-19-related negative affect symptoms. *Translational Psychiatry* **14**.

Park, H.-J. and Friston, K. (2013). Structural and functional brain networks: From connections to cognition. *Science* **342** 1238411.

Pascal, L. E., True, L. D., Campbell, D. S., Deutsch, E. W., Risk, M., Coleman, I. M., Eichner, L. J., Nelson, P. S. and Liu, A. Y. (2008). Correlation of mrna and protein levels: cell type-specific gene expression of cluster designation antigens in the prostate. *BMC genomics* **9** 1–13.

Pearson, K. (1901). Liii. on lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin philosophical magazine and journal of science* **2** 559–572.

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience* **9** 148–158.

Plis, S. M., Sui, J., Lane, T., Roy, S., Clark, V. P., Potluru, V. K., Huster, R. J., Michael, A., Sponheim, S. R., Weisend, M. P. and Calhoun, V. D. (2014). High-order interactions observed in multi-task intrinsic networks are dominant indicators of aberrant brain function in schizophrenia. *NeuroImage* **102** 35–48.

Pope, M., Fukushima, M., Betzel, R. F. and Sporns, O. (2021). Modular origins of high-amplitude cofluctuations in fine-scale functional connectivity dynamics. *Proceedings of the National Academy of Sciences* **118** e2109380118.

Popovic, D., Ruef, A., Dwyer, D. B., Antonucci, L. A., Eder, J., Sanfelici, R., Kambeitz-Ilankovic, L., Oztuerk, O. F., Dong, M. S., Paul, R. et al. (2020). Traces of trauma: a multivariate pattern analysis of childhood trauma, brain structure, and clinical phenotypes. *Biological Psychiatry* **88** 829–842.

Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., Vogel, A. C., Laumann, T. O., Miezin, F. M., Schlaggar, B. L. and Petersen, S. E. (2011a). Functional network organization of the human brain. *Neuron* **72** 665–678.

Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., Vogel, A. C., Laumann, T. O., Miezin, F. M., Schlaggar, B. L. et al. (2011b). Functional network organization of the human brain. *Neuron* **72** 665–678.

Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L. and Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fmri. *NeuroImage* **84** 320–341.

Pulvermüller, F. and Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience* **11** 351–360.

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G. and Sutskever, I. (2021). Learning transferable visual models from natural language supervision. In *International Conference on Machine Learning*.

Rahaman, N., Baratin, A., Arpit, D., Draxler, F., Lin, M., Hamprecht, F., Bengio, Y. and Courville, A. (2019). On the spectral bias of neural networks. In *Proceedings of the 36th International Conference on Machine Learning* (K. Chaudhuri and R. Salakhutdinov, eds.), vol. 97 of *Proceedings of Machine Learning Research*. PMLR.

Ravikumar, P., Wainwright, M. J., Raskutti, G. and Yu, B. (2011). High-dimensional covariance estimation by minimizing $\ell 1$-penalized log-determinant divergence. *Electronic Journal of Statistics* **5** 935–980.

Reggio, G. and Coppola, F. (1982). Koyaanisqatsi [motion picture]. *United States: Institute for Regional Education/American Zoetrope* .

Rishel, C. A., Huang, G. and Freedman, D. J. (2013). Independent category and spatial encoding in parietal cortex. *Neuron* **77** 969–979.

Roberts, J. and Robinson, P. (2012). Quantitative theory of driven nonlinear brain dynamics. *NeuroImage* **62** 1947–1955.

Rosenberg, M. D., Casey, B. and Holmes, A. J. (2018). Prediction complements explanation in understanding the developing brain. *Nature communications* **9** 589.

Rosenberg, M. D. and Finn, E. S. (2022). How to establish robust brain–behavior relationships without thousands of individuals. *Nature Neuroscience* **25** 835–837.

Rothman, A. J., Levina, E. and Zhu, J. (2009). Generalized thresholding of large covariance matrices. *Journal of the American Statistical Association* **104** 177–186.

Salvador, R., Suckling, J., Coleman, M. R., Pickard, J. D., Menon, D. and Bullmore, E. (2005). Neurophysiological architecture of functional magnetic resonance images of human brain. *Cerebral cortex* **15** 1332–1342.

Satpute, A. B. and Lindquist, K. A. (2019). The default mode network’s role in discrete emotion. *Trends in Cognitive Sciences* **23** 851–864.

Schaefer, A., Kong, R., Gordon, E., Laumann, T., Zuo, X.-N., Holmes, A., Eickhoff, S. and Yeo, B. T. (2018). Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cerebral Cortex* **28** 3095–3114.

Schneidman, E., Berry, M. J., Segev, R. and Bialek, W. (2006). Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* **440** 1007–1012.

Simony, E., Honey, C. J., Chen, J., Lositsky, O., Yeshurun, Y., Wiesel, A. and Hasson, U. (2016). Dynamic reconfiguration of the default mode network during narrative comprehension. *Nature Communications* **7** 12141.

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

Smith, S. M., Nichols, T. E., Vidaurre, D., Winkler, A. M., Behrens, T. E. J., Glasser, M. F., Uğurbil, K., Barch, D. M., Essen, D. C. V. and Miller, K. L. (2015). A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nature neuroscience* **18** 1565 – 1567.

Snoek, L., van der Miesen, M. M., Beemsterboer, T., van der Leij, A., Eigenhuis, A. and Scholte, H. S. (2020). The amsterdam open mri collection, a set of multimodal mri datasets for individual difference analyses. *bioRxiv* .

Sonkusare, S., Breakspear, M. and Guo, C. (2019). Naturalistic stimuli in neuroscience: Critically acclaimed. *Trends in Cognitive Sciences* **23** 699–714.

Spiegler, A., Knösche, T. R., Schwab, K., Haueisen, J. and Atay, F. M. (2011). Modeling brain resonance phenomena using a neural mass model. *PLOS Computational Biology* **7** 1–17.

Spielberger, C. D., Gonzalez-Reigosa, F., Martinez-Urrutia, A., Natalicio, L. F. and Natalicio, D. S. (1971). The state-trait anxiety inventory. *Revista Interamericana de Psicologia/Interamerican journal of psychology* **5**.

Stam, C. (2004). Functional connectivity patterns of human magnetoencephalographic recordings: a 'small-world' network? *Neuroscience Letters* **355** 25–28.

Stephan, K. E., Kasper, L., Harrison, L. M., Daunizeau, J., den Ouden, H. E., Breakspear, M. and Friston, K. J. (2008). Nonlinear dynamic causal models for fmri. *Neuroimage* **42** 649–662.

Sun, A., Wang, J. and Zhang, J. (2023). Identifying autism spectrum disorder using edge-centric functional connectivity. *Cerebral Cortex* **33** 8122–8130.

Sun, W. W. and Li, L. (2017). Store: sparse tensor response regression and neuroimaging analysis. *Journal of Machine Learning Research* **18** 1–37.

Supekar, K., Menon, V., Rubin, D., Musen, M. and Greicius, M. D. (2008). Network analysis of intrinsic functional brain connectivity in alzheimer's disease. *PLOS Computational Biology* **4** 1–11.

Székely, G. J., Rizzo, M. L. and Bakirov, N. K. (2007). Measuring and testing dependence by correlation of distances. *The Annals of Statistics* **35** 2769 – 2794.

Tan, K. M., Lu, J., Zhang, T. and Liu, H. (2021). Estimating and inferring the maximum degree of stimulus-locked time-varying brain connectivity networks. *Biometrics* **77** 379–390.

Tancik, M., Srinivasan, P. P., Mildenhall, B., Fridovich-Keil, S., Raghavan, N., Singhal, U., Ramamoorthi, R., Barron, J. T. and Ng, R. (2020). Fourier features let networks learn high frequency functions in low dimensional domains. *NeurIPS* .

van den Heuvel, M. P. and Sporns, O. (2013). An anatomical substrate for integration among functional networks in human cortex. *Journal of Neuroscience* **33** 14489–14500.

Van Den Heuvel, M. P., Stam, C. J., Kahn, R. S. and Pol, H. E. H. (2009). Efficiency of functional brain networks and intellectual performance. *Journal of Neuroscience* **29** 7619–7624.

Van der Ploeg, H. (1980). Validity of the zelf-beoordelings-vragenlijst (a dutch version of the spielberger state-trait anxiety inventory). *Nederlands Tijdschrift voor de Psychologie en haar Grensgebieden* .

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K. and WU-Minn HCP Consortium (2013). The WU-Minn Human Connectome Project: an overview. *Neuroimage* **80** 62–79.

Varangis, E., Habeck, C. G., Razlighi, Q. R. and Stern, Y. (2019). The effect of aging on resting state connectivity of predefined networks in the brain. *Frontiers in aging neuroscience* **11** 234.

Varley, T., Pope, M., Faskowitz, J. and Sporns, O. (2023). Multivariate information theory uncovers synergistic subsystems of the human cerebral cortex. *Communications Biology* **6**.

Váša, F. and Mišić, B. (2022). Null models in network neuroscience. *Nature Reviews Neuroscience* **23** 493–504.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. u. and Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan and R. Garnett, eds.), vol. 30. Curran Associates, Inc.

Ville, D. V. D., Farouj, Y., Preti, M. G., Liégeois, R. and Amico, E. (2021). When makes you unique: Temporality of the human brain fingerprint. *Science Advances* **7** eabj0751.

Vincent, J. L., Kahn, I., Snyder, A. Z., Raichle, M. E. and Buckner, R. L. (2008). Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *Journal of Neurophysiology* **100** 3328–3342. PMID: 18799601.

Viviani, R. (2013). Emotion regulation, attention to emotion, and the ventral attentional network. *Frontiers in Human Neuroscience* **Volume 7 - 2013**.

Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., Rosen, B. R. and Buckner, R. L. (1998). Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. *Science* **281** 1188–1191.

Wang, L., LaViolette, P., O'Keefe, K., Putcha, D., Bakkour, A., Van Dijk, K. R., Pihlajamäki, M., Dickerson, B. C. and Sperling, R. A. (2010). Intrinsic connectivity between the hippocampus and posteromedial cortex predicts memory performance in cognitively intact older individuals. *Neuroimage* **51** 910–917.

Wang, Z., Williams, V. J., Stephens, K. A., Kim, C.-M., Bai, L., Zhang, M. and Salat, D. H. (2020). The effect of white matter signal abnormalities on default mode network connectivity in mild cognitive impairment. *Human Brain Mapping* **41** 1237–1248.

Wen, H., Liu, Y., Rekik, I., Wang, S., Chen, Z., Zhang, J., Zhang, Y., Peng, Y. and He, H. (2018). Combining disrupted and discriminative topological properties of functional connectivity networks as neuroimaging biomarkers for accurate diagnosis of early tourette syndrome children. *Molecular neurobiology* **55** 3251–3269.

Wen, Z., Raio, C. M., Pace-Schott, E. F., Lazar, S. W., LeDoux, J. E., Phelps, E. A. and Milad, M. R. (2022). Temporally and anatomically specific contributions of the human amygdala to threat and safety learning. *Proceedings of the National Academy of Sciences* **119** e2204066119.

Weng, S.-J., Wiggins, J. L., Peltier, S. J., Carrasco, M., Risi, S., Lord, C. and Monk, C. S. (2010). Alterations of resting state functional connectivity in the default network in adolescents with autism spectrum disorders. *Brain Research* **1313** 202–214.

Wold, H. (1975). Soft modelling by latent variables: The non-linear iterative partial least squares (nipals) approach. *Journal of Applied Probability* **12** 117–142.

Wold, S., Ruhe, A., Wold, H. and Dunn, W. J., III (1984). The collinearity problem in linear regression. the partial least squares (pls) approach to generalized inverses. *SIAM Journal on Scientific and Statistical Computing* **5** 735–743.

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M. et al. (2020). Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations*.

Woo, C.-W., Chang, L. J., Lindquist, M. A. and Wager, T. D. (2017). Building better biomarkers: brain models in translational neuroimaging. *Nature Neuroscience* **20** 365–377.

Xu, Y. and Lindquist, M. (2015). Dynamic connectivity detection: an algorithm for determining functional connectivity change points in fMRI data. *Frontiers in Neuroscience* **9** 285.

Yahata, N., Morimoto, J., Hashimoto, R., Lisi, G., Shibata, K., Kawakubo, Y., Kuwabara, H., Kuroda, M., Yamada, T., Megumi, F. et al. (2016). A small number of abnormal brain connections predicts adult autism spectrum disorder. *Nature communications* **7** 11254.

Yan, C.-G., Chen, X., Li, L., Castellanos, F. X., Bai, T.-J., Bo, Q.-J., Cao, J., Chen, G.-M., Chen, N.-X., Chen, W., Cheng, C., Cheng, Y.-Q., Cui, X.-L., Duan, J., Fang, Y.-R., Gong, Q.-Y., Guo, W.-B., Hou, Z.-H., Hu, L., Kuang, L., Li, F., Li, K.-M., Li, T., Liu, Y.-S., Liu, Z.-N., Long, Y.-C., Luo, Q.-H., Meng, H.-Q., Peng, D.-H., Qiu, H.-T., Qiu, J., Shen, Y.-D., Shi, Y.-S., Wang, C.-Y., Wang, F., Wang, K., Wang, L., Wang, X., Wang, Y., Wu, X.-P., Wu, X.-R., Xie, C.-M., Xie, G.-R., Xie, H.-Y., Xie, P., Xu, X.-F., Yang, H., Yang, J., Yao, J.-S., Yao, S.-Q., Yin, Y.-Y., Yuan, Y.-G., Zhang, A.-X., Zhang, H., Zhang, K.-R., Zhang, L., Zhang, Z.-J., Zhou, R.-B., Zhou, Y.-T., Zhu, J.-J., Zou, C.-J., Si, T.-M., Zuo, X.-N., Zhao, J.-P. and Zang, Y.-F. (2019). Reduced default mode network functional connectivity in patients with recurrent major depressive disorder. *Proceedings of the National Academy of Sciences* **116** 9078–9083.

Yang, B., Wang, M., Zhou, W., Wang, X., Chen, S., Yuan, L.-X. and Dong, G.-H. (2023). Edge-centric functional network analyses reveal disrupted network configuration in autism spectrum disorder. *Journal of Affective Disorders* **336** 74–80.

Yang, Y.-L., Deng, H.-X., Xing, G.-Y., Xia, X.-L. and Li, H.-F. (2015). Brain functional network connectivity based on a visual task: visual information processing-related brain regions are significantly activated in the task state. *Neural Regeneration Research* **10** 298–307.

Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R. and Le, Q. V. (2019). *XLNet: generalized autoregressive pretraining for language understanding*. Curran Associates Inc., Red Hook, NY, USA.

Yeo, B. T., Krienen, F. M., Chee, M. W. and Buckner, R. L. (2014). Estimates of segregation and overlap of functional connectivity networks in the human cerebral cortex. *NeuroImage* **88** 212–227.

Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R. et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of neurophysiology* .

Yi, F. and Zou, H. (2013). Sure-tuned tapering estimation of large covariance matrices. *Computational Statistics & Data Analysis* **58** 339–351. The Third Special Issue on Statistical Signal Extraction and Filtering.

Zamani Esfahlani, F., Byrge, L., Tanner, J., Sporns, O., Kennedy, D. P. and Betzel, R. F. (2022). Edge-centric analysis of time-varying functional brain networks with applications in autism spectrum disorder. *NeuroImage* **263** 119591.

Zanto, T. P. and Gazzaley, A. (2013). Fronto-parietal network: flexible hub of cognitive control. *Trends in Cognitive Sciences* **17** 602–603.

Zatorre, R. J., Fields, R. D. and Johansen-Berg, H. (2012). Plasticity in gray and white: neuroimaging changes in brain structure during learning. *Nature neuroscience* **15** 528–536.

Zhang, D., Li, L., Sripada, C. and Kang, J. (2023). Image response regression via deep neural networks. *Journal of the Royal Statistical Society Series B: Statistical Methodology* **85** 1589–1614.

Zhang, Y., Zhang, H., Chen, X., Lee, S.-W. and Shen, D. (2017). Hybrid High-order Functional Connectivity Networks Using Resting-state Functional MRI for Mild Cognitive Impairment Diagnosis. *Scientific Reports* **7** 6530.

Zhou, L., Zhou, Y., Corso, J. J., Socher, R. and Xiong, C. (2018). End-to-end dense video captioning with masked transformer. In *Proceedings of the IEEE conference on computer vision and pattern recognition*.

Zhu, H., Fan, J. and Kong, L. (2014). Spatially varying coefficient model for neuroimaging data with jump discontinuities. *Journal of the American Statistical Association* **109** 1084–1098. PMID: 25435598.