

**Multimodal Human Brain Connectivity Analysis based on
Graph Theory**

by

Chendi Wang

B.A.Sc, Biomedical Engineering, Southeast University, 2010

M.A.Sc, Biomedical Engineering, Southeast University, 2013

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL
STUDIES

(Electrical and Computer Engineering)

The University of British Columbia

(Vancouver)

April 2018

© Chendi Wang, 2018

Abstract

Billions of people worldwide are affected by neurological disorders. Recent studies indicate that many neurological disorders can be described as dysconnectivity syndromes, and associated with changes in the brain networks prior to the development of clinical symptoms.

This thesis presents contributions towards improving brain connectivity analysis based on graph theory representation of the human brain network. We propose novel multimodal techniques to analyze brain imaging data to better understand its structure, function and connectivity, i.e., brain connectomics.

Our first contribution is towards improving parcellation, i.e., brain network node definition, in terms of reproducibility, functional homogeneity, leftout data likelihood and overlaps with cytoarchitecture, by utilizing the neighbourhood information and multi-modality integration techniques. Specifically, we embed neighbourhood connectivity information into the affinity matrix for parcellation to ameliorate the adverse effects of noise. We further integrate the connectivity information from both anatomical and functional modalities based on adaptive weighting for an improved parcellation.

Our second contribution is to propose noise reduction techniques for brain network edge definition. We propose a matrix completion based technique to combat false negatives by recovering missing connections. We also present a local thresholding method which can address the regional bias issue when suppressing the false positives in connectivity estimates.

Our third contribution is to improve the brain subnetwork extraction by using multi-pronged graphical metric guided methods. We propose a connection-fingerprint based modularity reinforcement model which reflects the putative mod-

ular structure of a brain graph. Inspired by the brain subnetwork's biological nature, we propose a provincial hub guided feedback optimization model for more reproducible subnetwork extraction.

Our fourth contribution is to develop multimodal integration techniques to further improve brain subnetwork extraction. We propose a provincial hub guided subnetwork extraction model to fuse anatomical and functional data by propagating the modular structure information across different modalities. We further propose to fuse the task and rest functional data based on hypergraphs for non-overlapping and overlapping subnetwork extraction.

Our results collectively indicate that combing multimodal information and applying graphical metric guided strategies outperform classical unimodal brain connectivity analysis methods. The resulting methods could provide important insights into cognitive and clinical neuroscience.

Lay Summary

3.6 million people in Canada (more than 10% of the total population) are affected by neurological diseases, such as Parkinson's disease, Alzheimer's disease and other dementias. There are changes in the anatomy of the human brain and changes in the way how brain works before a patient develops clinical symptoms. In this thesis, we present new methods to study human brain images to better understand its anatomy and how different parts of the brain work together to execute certain functions. Using multiple sources of brain images, we model the brain as a graph, by defining different parts of the brain as nodes and identifying the relations between those parts. The relations between brain parts are studied by how strongly they are connected to each other, and how they work together. These methods will improve our understanding of how human brain works and furthermore how the brain is affected by diseases.

Preface

The data used in research presented herein was approved by the Ethics Board from the Open Access Dataset.

This thesis is based on the following papers, resulting from collaboration between multiple researches.

Studies described in Chapter 2 have been published in:

[P1] [1] (ISBI'15) C. Wang, B. Yoldemir, and R. Abugharbieh. Improved functional cortical parcellation using a neighborhood-information-embedded affinity matrix. In *IEEE International Symposium on Biomedical Imaging*, pages 1340–1343, 2015.

[P2] [2] (MICCAI'15) C. Wang, B. Yoldemir, and R. Abugharbieh. Multimodal cortical parcellation based on anatomical and functional brain connectivity. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 21–28. Springer, 2015. Poster presentation - Acceptance rate: $\sim 33\%$. Early Acceptance. Student travel award.

[P3] [3] (Journal'18) C. Wang, B. Ng, and R. Abugharbieh. Multimodal whole brain parcellation, submitted, xx(x):xxxx–xxxx, 2018.

Studies described in Chapter 3 have been published in:

[P4] [4] (MICCAI'17 Workshop) C. Wang, B. Ng, A. Amir-Khalili, and R. Abugharbieh. Recovering missing connections in diffusion weighted mri using matrix completion. In *International Conference on Medical Image Computing and Computer-Assisted Intervention Workshop on Computational Diffusion MRI*. Springer, 2017.

[P5] [5] (MICCAI'16) C. Wang, B. Ng, and R. Abugharbieh. Modularity reinforcement for improving brain subnetwork extraction. In *International Conference*

on *Medical Image Computing and Computer-Assisted Intervention*, pages 132–139. Springer, 2016. Poster presentation - Acceptance rate: $\sim 30\%$. Student travel award.

Studies described in Chapter 4 have been published in:

[P5] [5] (MICCAI'16) C. Wang, B. Ng, and R. Abugharbieh. Modularity reinforcement for improving brain subnetwork extraction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 132–139. Springer, 2016. Poster presentation - Acceptance rate: $\sim 30\%$. Student travel award.

[P6] [6] (IPMI'17) C. Wang, B. Ng, and R. Abugharbieh. Multimodal brain subnetwork extraction using provincial hub guided random walks. In *International Conference on Information Processing in Medical Imaging*, pages 287–298. Springer, Cham, 2017. Oral presentation - Acceptance rate: $\sim 15\%$. Scholarship for Junior Scientists for Underrepresented Populations.

Studies described in Chapter 5 have been published in:

[P6] [6] (IPMI'17) C. Wang, B. Ng, and R. Abugharbieh. Multimodal brain subnetwork extraction using provincial hub guided random walks. In *International Conference on Information Processing in Medical Imaging*, pages 287–298. Springer, Cham, 2017. Oral presentation - Acceptance rate: $\sim 15\%$. Scholarship for Junior Scientists for Underrepresented Populations.

[P7] [7] (Journal'18) C. Wang and R. Abugharbieh. High order relation informed task and rest data integration for subnetwork extraction based on hypergraph, submitted, xx(x):xxxx–xxxx, 2018.

[P8] [8] (Journal'18) C. Wang and R. Abugharbieh. Co-activated clique based multisource overlapping subnetwork extraction, submitted, xx(x):xxxx–xxxx, 2018.

All listed publications were revised and edited by all co-authors.

The contribution statements are as the following:

In all the listed publications, Wang *et al.* [1–8], the primary author, I was the main contributor to the algorithmic idea conception, design, implementation, and testing of proposed methodology, and the writing effort under the supervision of Dr. Rafeef Abugharbieh. I also presented the oral presentation for Wang *et al.* [6] and the poster presentation for Wang *et al.* [1, 2, 5].

In Wang *et al.* [1] and Wang *et al.* [2], I was responsible for concept formation

and design for the idea, data downloading, processing and analysis, experimental design, as well as manuscript composition. Dr. Burak Yoldemir under the supervision of Dr. Abugharbieh helped with preprocessing of the data and provided suggestions on the manuscript revision.

In Wang *et al.*. [3, 5], I was responsible for concept formation and design for the idea, data preprocessing and analysis, and manuscript composition. Dr. Bernard Ng helped immensely with his valuable input on experimental design and manuscript edits.

In Wang *et al.*. [6], I was responsible for concept formation and design, data preprocessing and analysis, and manuscript composition. Dr. Bernard Ng helped with the ideas for experimental design.

In Wang *et al.*. [4], Dr. Bernard Ng contributed the general matrix completion idea and helped with brainstorming the validation schemes, Dr. Alborz Amir-Khalili under the supervision of Dr. Abugharbieh contributed the rank range selection idea, and I contributed the neighborhood information filling idea. I was responsible for processing and analysis of the data, implementation of the methodology, validation experimental design, and manuscript composition. Both Dr. Bernard Ng and Dr. Alborz Amir-Khalili contributed to manuscript edits.

In [7] and [8], I was responsible for concept formation and design, data preprocessing and analysis, experimental design, and manuscript composition.

Table of Contents

Abstract	ii
Lay Summary	iv
Preface	v
Table of Contents	viii
List of Tables	xvii
List of Figures	xviii
List of Acronyms	xxx
Acknowledgments	xxxv
1 Introduction	1
1.1 Motivation	1
1.2 Thesis Objectives	4
1.3 Imaging for Brain Connectomics	4
1.3.1 Definition of Brain Connectomics	4
1.3.2 Imaging Resolution in Brain Connectomics	5
1.3.3 Functional Imaging for Macroscale Connectomics	5
1.3.3.1 Modalities for Functional Imaging	5
1.3.3.2 Functional Magnetic Resonance Imaging (fMRI) Experimental Design	7

1.3.4	Structural Imaging for Macroscale Connectomics . . .	7
1.4	State-of-the-art Techniques in Brain Connectivity Analysis . .	8
1.4.1	Brain Network Analysis for Connectivity Studies . . .	8
1.4.1.1	Model-dependent Methods for Seed-based Network Analysis	8
1.4.1.2	Model-free Methods for Whole Brain Network Analysis	9
1.4.2	Graph Representation in Brain Connectomics	10
1.4.2.1	Existing Studies of Graph Use for Brain Connectomics	10
1.4.2.2	Definition of Nodes: Parcellation-based Brain Connectome	11
1.4.2.3	Definition of Edges: Estimating Connections Between Brain Regions	12
1.4.2.3.1	Estimating Anatomical Connectivity	13
1.4.2.3.2	Estimating Functional Connectivity	17
1.4.2.3.3	Thresholding	19
1.4.2.4	Network Measures: Graphical Metrics for Brain Connectomics	19
1.4.2.5	Module Detection: Brain Subnetwork Extraction	20
1.4.2.5.1	Anatomical Subnetworks	22
1.4.2.5.2	Functional Subnetworks	22
1.4.3	Multimodal Fusion in Brain Connectivity Analysis . .	25
1.4.4	Deep Learning in Brain Connectivity Analysis	26
1.5	Current Challenges Addressed in This Thesis	28
1.6	Problem Statement	29
1.6.1	How Can We Improve Parcellation?	30
1.6.2	How Can We Achieve Noise Reduction When Constructing Edges?	30
1.6.3	How Can We Improve Subnetwork Extraction?	30
1.6.4	How Can Multimodal Information Help with Subnetwork Extractions for Brain Connectivity Analysis?	30

1.7	Thesis Contributions	30
1.7.1	Brain Network Node Definition	31
1.7.2	Brain Network Edge Estimation	31
1.7.3	Graphical Metric Guided Subnetwork Extraction	32
1.7.4	Multimodal Fusion for Subnetwork Extraction	32
1.8	Materials and Experimental Setup	33
1.8.1	Overview of Currently Available Open Access Neuroimaging Datasets	33
1.8.2	Human Connectome Project	33
1.8.3	Data Preprocessing	34
1.8.3.1	Data Preprocessing for Functional Connectivity	34
1.8.3.2	Data Preprocessing for Anatomical Connectivity	35
2	Neighbourhood Information Embedding and Multimodal Integration for Improved Parcellation	36
2.1	Neighborhood Connectivity Informed Parcellation	37
2.1.1	Related Work	37
2.1.2	Neighborhood Information Embedded Affinity Matrix	37
2.1.2.1	Gaussian Kernel Affinity Matrix	38
2.1.2.2	Proposed Affinity Matrix	38
2.1.3	Experiments and Results	40
2.1.4	Discussion	43
2.2	Multimodal Connectivity Fusion for Parcellation	44
2.2.1	Related Work	44
2.2.2	Proposed Multimodal Integrated Parcellation Framework	44
2.2.3	Anatomical and Functional Connectivity Estimation	46
2.2.4	Distribution Normalization	47
2.2.5	Multimodal Connectivity Estimation using Adaptive Weighting	48
2.2.6	Clustering and Region Level Extension	49
2.2.7	Design of Evaluation Metrics	50
2.2.7.1	Reproducibility	50

2.2.7.2	Functional Homogeneity	51
2.2.7.3	Leftout Data Likelihood	51
2.2.7.4	Overlaps with Cytoarchitecture	52
2.2.7.5	Subnetwork Extraction	52
2.2.8	Experiments and Results	53
2.2.8.1	Reproducibility	53
2.2.8.2	Homogeneity	57
2.2.8.3	Leftout Data Likelihood	58
2.2.8.4	Cytoarchitecture	61
2.2.8.5	Subnetwork Structure	63
2.2.9	Discussion	65
2.2.9.1	Purposes and Key Challenges of Multimodal Parcellation	65
2.2.9.2	Multimodal Parcellation Improves Reproducibility and Data Likelihood, and Maintains Functional Homogeneity	65
2.2.9.3	Linking Parcels to Prior Knowledge	66
2.2.9.4	Region Level vs. Whole-brain Parcellation	66
2.3	Summary	67
3	Noise Reduction for Brain Network Edge Building	69
3.1	Matrix Completion to Combat False Negatives	69
3.1.1	Related Work	69
3.1.2	Low Rank Matrix Completion for Connectivity Estimation	70
3.1.3	Rank Range Search and Aggregation	71
3.1.4	Negative Entries Filling using Neighborhood Information	72
3.1.5	Experiments	73
3.1.5.1	Materials	73
3.1.5.1.1	Synthetic Data	73
3.1.5.1.2	Real Data	73
3.1.5.2	Results	75
3.1.5.2.1	Recovery Accuracy	76

3.1.5.2.2	Intelligence Quotient (IQ) Prediction	77
3.1.6	Discussion	81
3.2	Local Thresholding to Suppress False Positives	82
3.2.1	Related Work	82
3.2.2	Local Thresholding	84
3.2.3	Experiments	85
3.2.3.1	Materials	85
3.2.3.1.1	Synthetic Data	85
3.2.3.1.2	Real Data	85
3.2.3.2	Results	87
3.2.3.2.1	Synthetic Data	87
3.2.3.2.2	Real Data	88
3.2.4	Discussion	90
3.3	Summary	90
4	Graphical Metric Guided Subnetwork Extraction	92
4.1	Modularity Reinforcement Subnetwork Extraction	93
4.1.1	Related Work	93
4.1.2	Modularity Reinforcement Model	93
4.1.3	Subnetwork Extraction Based on Graph Cuts	94
4.1.4	Experiments	94
4.1.4.1	Materials	94
4.1.4.2	Results	94
4.1.4.2.1	Synthetic Data	95
4.1.4.2.2	Real Data	96
4.1.5	Discussion	98
4.2	Provincial Hub Guided Random Walker Based Subnetwork Ex- traction	100
4.2.1	Related Work	100
4.2.2	Resemblance Between Provincial Hubs and the Seeds	101
4.2.3	Overview of Random Walker Model	102
4.2.4	Feedback Informed Optimization Model	103
4.2.4.1	Relating to Random Walker (RW) with Prior Model	103

4.2.4.2	Multi-seed Model	104
4.2.4.3	Subnetwork Assignment Confidence Based on Posterior Probability	105
4.2.4.4	Framework of the Feedback Informed Optima- tion Model	106
4.2.5	Experiments	107
4.2.5.1	Materials	107
4.2.5.1.1	Synthetic Data	107
4.2.5.1.2	Real Data	108
4.2.5.1.3	Parameter Setting	108
4.2.5.2	Results	108
4.2.5.2.1	Synthetic Data	108
4.2.5.2.2	Real Data	109
4.2.6	Discussion	109
4.3	Summary	110
5	Multimodal/Multisource Brain Subnetwork Extraction	111
5.1	Multimodal Random Walker based Subnetwork Extraction . .	112
5.1.1	Related Work	112
5.1.2	Multimodal Provincial Hub Guided Random Walker Model	113
5.1.3	Subnetwork Assignment Confidence and Overlapping Sub- network Exploration Based on Posterior Probability . .	115
5.1.4	Experiments	115
5.1.4.1	Materials	115
5.1.4.1.1	Synthetic Data	115
5.1.4.1.2	Real Data	115
5.1.4.2	Results	116
5.1.4.2.1	Synthetic Data	116
5.1.4.2.2	Real Data	117
5.1.5	Discussion	119
5.2	Fusion of Task and Resting State Functional Connectivity for Sub- network Extraction Based on Hypergraph	122

5.2.1	Related Work - Relationship between Task and Resting Functional Connectivity	122
5.2.2	Related Work - Multilayer Brain Network Analysis . .	123
5.2.3	High Order Relation Informed Subnetwork Extraction	123
5.2.3.1	Framework	124
5.2.3.2	Notation Overview of Hypergraph	124
5.2.3.2.1	Notations	124
5.2.3.2.2	Graphcut of the Hypergraph	126
5.2.3.3	Task Activation Detection - Node Definition in the Hypergraph	127
5.2.3.4	Strength Informed Weighted Multi-task Hypergraph	128
5.2.3.4.1	Pairwise Nodal Connection Strength Estimation	128
5.2.3.4.2	Proposed Strength Informed Weighted Hypergraph	129
5.2.3.5	Multisource Integration of Rest and Task fMRI	130
5.2.4	Experiments	131
5.2.4.1	Materials	132
5.2.4.2	Similarity of Functional Connectivity (FC) Between Resting State and Task Data	133
5.2.4.3	Modularity Q Value	134
5.2.4.4	Inter-subject Reproducibility of Subnetwork Extraction	135
5.2.4.5	Biological Meaning	136
5.2.5	Discussion	141
5.2.5.1	Hypergraph Encodes Higher Order Nodal Relationship	141
5.2.5.2	Multisource Integration Improves Subnetwork Extraction	141
5.2.5.3	Limitations and Future Directions	142
5.3	Clique Based Multisource Overlapping Brain Subnetwork Extraction	142

5.3.1	Related Work - Overlapping Brain Subnetwork Extraction	142
5.3.2	Co-activated Clique Based Multisource Overlapping Subnetwork Extraction	144
5.3.2.1	Clique Identification Based on Task Co-activation	146
5.3.2.2	Clique Property Computation	149
5.3.2.3	Core Clique Identification	150
5.3.2.4	Clique Based Overlapping Subnetwork Extraction	150
5.3.3	Experiments	151
5.3.3.1	Materials	152
5.3.3.2	Comparison with Existing Overlapping Subnetwork Extraction Methods	153
5.3.3.2.1	Group-level Subnetwork Extraction Reproducibility	153
5.3.3.2.2	Subject-wise Subnetwork Extraction Reproducibility	154
5.3.3.3	Biological Meaning - Analyzing Function Integration	154
5.3.3.4	Comparison Between the Subnetwork Overlaps and the RW Posterior Probability	157
5.3.4	Discussion	159
5.3.4.1	Benefits of Clique Identification Based on Task Co-activation	159
5.3.4.2	Multisource Information Integration Improves the Overlapping Subnetwork Extraction	160
5.3.4.3	Overlapping Subnetwork Identification Corresponds with the RW Posterior Probability	160
5.3.4.4	Other Considerations	161
5.3.4.5	Limitations and Future Work	161
5.4	Summary	161
6	Conclusions	163

6.1	Our Contributions in Graphical Framework for Human Brain Connectivity Analysis	163
6.1.1	Definition of Nodes: Parcellation-based Brain Connectome	163
6.1.2	Definition of Edges: Estimating Connections Between Brain Regions	164
6.1.3	Network Measures: Graphical Metrics for Brain Connectomics	165
6.1.4	Multimodal/Multisource Fusion to Improve Brain Connectivity Analysis	166
6.2	Future Work	168
	Bibliography	171

List of Tables

Table 2.1	Details on existing anatomical and functional parcellations.	61
-----------	--	----

List of Figures

Figure 1.1	Prevalence and total Disability-Adjusted Life Years (DALYs) for selected diseases and neurological disorders. Data are adopted from [9, 10].	2
Figure 1.2	Projected indirect economic costs due to working-age death and disability, by select neurological condition and age group, Canada, 2011, 2021, and 2031	3
Figure 1.3	Spatial-temporal resolution plots for commonly used functional neuroimaging methods in the year of 2014. VSD, voltage-sensitive dye; TMS, transcranial magnetic stimulation; 2-DG, 2-deoxyglucose. Image courtesy of [11].	6
Figure 1.4	Connectivity matrix based on graph theory created by Rolf Kötter and projections derived from tract tracing, one of the first attempts to obtain a large-scale anatomical connectome of the mammalian brain. Image courtesy of [12].	11
Figure 1.5	Cartoon illustrations of ambiguities in mapping diffusion to axon geometry. Different axon geometries can lead to a similarly oriented tensor. The tensor's principal direction is the same for all cases, but modeling multiple Principal Diffusion Direction (PDD)s helps distinguish a few of the cases. Modeling fiber fanning separates the top two cases. Further modeling the polarity of a fanning can help separate all cases. Image courtesy of [13].	14

Figure 1.6	Deterministic and probabilistic tractography. Pyramidal tract streamlines based on deterministic (left) and probabilistic (right) approaches (results superimposed on Fractional Anisotropy (FA) image). Color bar indicates confidence about the presence of the tract. Image courtesy of [14].	15
Figure 1.7	Typical workflow for Anatomical Connectivity (AC) estimation based on a Diffusion-weighted Magnetic Resonance Imaging (dMRI) experiment. Diffusion weighted images are acquired and preprocessed to reconstruct the diffusion Orientation Distribution Function (ODF)s. Fiber tracts are then generated using tractography. AC matrix can indicate the fiber counts between each brain pair. MR scanner image: A patient is loaded into an MRI machine, by Morsa Images/Getty Images, https://www.verywell.com/diffusion-weighted-mri-3146133 , retrieved on August 30, 2017.	16
Figure 1.8	Example of false negatives in <i>in vivo</i> tractography of the cerebellum. The dotted green lines indicate the incomplete reconstruction of the superior cerebellar tracts where streamlines stop before crossing to the contralateral side as expected from known post-mortem neuroanatomy (false negatives). Image courtesy of [15].	17
Figure 1.9	Example of ambiguities near the gray-white matter interface. Both axon configurations lead to the same diffusion profile and tracking results, but with different connectivities. Image courtesy of [13].	17
Figure 1.10	Typical workflow for FC estimation based on a fMRI experiment. Functional Magnetic Resonance Imaging (MRI) images acquired when the subjects are at task conditions or rest. Voxel and regional time courses are extracted after image preprocessing. FC matrix can indicate covariance between each brain pair. MR scanner image: A patient is loaded into an MRI machine, by Morsa Images/Getty Images, https://www.verywell.com/diffusion-weighted-mri-3146133 , retrieved on August 30, 2017.	18

Figure 1.11	Examples of the connectivity matrices after graph density thresholding. Image courtesy of [16].	19
Figure 1.12	Exemplar network parameter measures. Image courtesy of [17].	21
Figure 1.13	Subnetwork structure based on functional connectivity. Subgraphs from three thresholds are shown for the areal (spheres) and modified voxel-wise graphs (surfaces). Subnetworks such as default mode subnetwork, frontoparietal executive control subnetwork, and attention subnetwork comprise of spatially distributed brain regions. Image courtesy of [18].	23
Figure 1.14	The functional coactivation network based on meta-analysis of task-related fMRI studies has similar modularity to a functional connectivity network based on resting-state fMRI data. Co-activation and connectivity networks plotted in anatomical space. The edges are defined by the minimum spanning tree for illustrative purposes. The size of the nodes is proportional to their strength, and their color corresponds to module membership. Image courtesy of [19].	24
Figure 2.1	Illustration of Interior Point (IP) (red dot) and Boundary Point (BP) (green and blue dots). The neighbors of the latter have a heterogeneous cluster membership structure whereas all neighbors of the former belong to the same cluster.	39
Figure 2.2	Theoretical illustration of Neighborhood-information-embedded Multiple Density (NMD) affinity matrix. Left is traditional Gaussian kernel, and the right is NMD. The affinity distribution, shown in the red bars, expands to a wider range when NMD is used.	40

Figure 2.3	Inter-subject test-retest reliability quantified using Dice Similarity Coefficient (DSC) in Inferior Parietal Lobule (IPL) and Cingulate Cortex (Cg) for the Human Connectome Project (HCP) data. As reflected in the box plots, our method consistently outperforms the state-of-art method. The blue rectangle spans the first quartile to third quartile. The red line indicates the median and the black whiskers indicate the minimum and maximum.	42
Figure 2.4	Qualitative parcellation results for IPL and Cg using the HCP data. In each area - Top: Using the multiple density in (2.2); Bottom: Using our proposed affinity matrix NMD in (2.4). The 1 st column shows the group maps while the 2 nd and 3 rd columns are the subject-specific maps with the highest and lowest DSC with group maps, respectively. The significant differences between the subject-specific map having the lowest DSC (right) and the group map (left) are highlighted with arrows.	43
Figure 2.5	Flowchart of proposed approach for multimodal brain parcellation. AC and FC estimates are generated for each voxel to derive voxel-by-voxel connectivity matrices followed by a distribution normalization. The resulting matrices are integrated to generate multimodal connectivity estimates. Normalized cuts (Ncuts) is applied on the integrated similarity matrices, to generate whole-brain parcellation.	45
Figure 2.6	Histograms of FC and AC derived from the group data of 77 HCP subjects. The difference between the two distributions could potentially cause bias if the FC and AC are naively fused.	48

Figure 2.7	Group-wise parcellation and Subject-group consistency. Both the parcellations based on the proposed multimodal approach (left two) and unimodal FC based parcellations (right two) are displayed at 512-parcel scale in 2D space. Second and fourth column: Subject level parcellation having the lowest DSC with the group parcellation. The proposed multimodal parcellations attain higher subject-group consistency (a & b), compared to unimodal FC based parcellations (c & d), which we highlight certain areas using blue circles.	54
Figure 2.8	Intra-subject test-retest reliability using the DSC as the evaluation criterion. The number of parcels M was set to 256, 512 and 1024, respectively. The blue rectangle spans the first quartile to third quartile. The red line indicates the median and the black whiskers indicate the minimum and maximum. Multimodal parcellations achieved significantly higher test-retest reliability than those generated based on fMRI data alone.	55
Figure 2.9	The inter-subject test-retest reliability (range from low red to high green) of each brain region of the Harvard-Oxford atlas (only shown on sagittal slices X=88mm, 72mm and 56mm as exemplars). Regions, such as the Lingual Gyrus (LG), Lateral Occipital Cortex (LOC), Postcentral Gyrus (PG), Frontal Pole (FP), and Superior Frontal Gyrus (SFG) show relatively lower reproducibility.	55
Figure 2.10	Subject-group consistency. Comparison between the group parcellation and the subject parcellation of frontal pole having the lowest DSC with the group parcellations.	56

Figure 2.11	Stability with respect to the number of subjects in generating group parcellations. We compared the multimodal group parcellations generated from all 77 subjects against N_S subjects with N_S set between 5 and 75 at interval of 5 subjects. Stability was estimated using the Dice coefficient. The number of parcels was set to 512. The stabilities of the group parcellations as measured using the Dice coefficient plateaued at 0.96 after $N_S = 50$	57
Figure 2.12	Functional homogeneity comparisons between multimodal and unimodal parcellations. In order to show consistent gain in each case of comparisons, plots of Multimodal Connectivity (MC)-FC, MC-AC, MC-physical distance (DC) and MC-random parcellation (RM) are provided. Positive FC pc ratio and negative mean distance indicate higher functional homogeneity gained from multimodal parcellations. Number of parcels were set to 512. Multimodal parcels achieved significantly higher functional homogeneity in both FC pc ratio and mean distance, see text for details.	59
Figure 2.13	Existing anatomical and functional parcellations	59
Figure 2.14	Functional homogeneity comparison between multimodal parcellations and existing atlases. Number of parcels for multimodal parcellation was set to the exact number of parcels in the contrasted atlas. Contrasted atlases included Harvard-Oxford (HO) (112 parcels), Automated Anatomical Labeling (AAL) (116 parcels), Gordon's (333 parcels), Shen's (268 parcels), and Craddock (190 parcels). Multimodal approach achieved significantly higher homogeneity against the contrasted atlases except for Gordon's.	60

Figure 2.15	Leftout log-data likelihood comparison between multimodal and unimodal parcellation. Number of parcels was set to 256, 512 and 1024, respectively. (a) The likelihood of MC was higher than FC and DC, and significantly higher than AC at the scale of 256 parcels. (b) The likelihood of MC was higher than FC and AC at the scale of 512 parcels. (c) The likelihood of MC was significantly higher than FC, AC and DC at the scale of 1024 parcels. RM performs better due to the regular sampling of volume of interest from parcels with roughly equal size.	62
Figure 2.16	Comparison of the parcel boundaries to exemplar cytoarchitectonic areas. The areas in cytoarchitectonic map are bordered in black dotted line, and the parcels having at least 70% overlap with the cytoarchitectonic areas are displayed using solid color based on the ‘RGBYR’ colormap in Caret [20].	63
Figure 2.17	Extracted subnetworks. Subnetworks, such as Visual Network (light green in Occipital Cortex), Ventral Motor and Sensory system and Auditory system (orange), Dorsal Motor and Sensory system (spring green), and the Default Mode network (dark red) have been found in (b).	64
Figure 3.1	Eigenspectrum of an exemplar AC	72
Figure 3.2	Matrix recovery on an exemplar synthetic dataset. MCmedFill (i) more accurately recovered the ground truth connections (in red) than the contrasted methods (c-h).	74
Figure 3.3	Matrix recovery using MCmedFill (a) on an exemplar synthetic dataset. The weights in each column of the low rank matrix \mathbf{Y} align with the ground truth subnetwork assignment. (b) An example of thresholded weights for MCmedFill shown.	75
Figure 3.4	Rank selection method. Our criteria of using recovery accuracy on randomly removed AC entries displays clear normalized root-mean-squared-error (NRMSE) minima for rank selection.	76

Figure 3.5	Matrix recovery on synthetic data. Mean values are indicated as black diamonds and labeled at the bottom of each subfigure. MCmedFill achieved significantly higher accuracy than contrasted methods based on Wilcoxon signed rank test.	78
Figure 3.6	Matrix recovery on real data. Mean values are indicated as black diamonds and labeled at the bottom of each subfigure. MCmedFill achieved significantly higher accuracy than contrasted methods based on Wilcoxon signed rank test.	79
Figure 3.7	IQ prediction on real data. (a) MCmedFill achieved significantly higher correlation coefficient between observed and predicted IQ scores than contrasted methods. (b) Brain region subnetwork weights along columns of \mathbf{Y} (thresholded for clearer visualization) were found to resemble known brain networks, e.g., Frontal-Parietal network.	80
Figure 3.8	Signal dropouts in the orbitofrontal cortex (OFC) caused by field inhomogeneities (at 1.5 T). Image courtesy of [21].	82
Figure 3.9	Schematic illustrating local thresholding (minimal spanning tree and k-nearest neighbors (MST-KNN)) and global thresholding methods. Image courtesy of [22].	83
Figure 3.10	Schematic illustrating our proposed local thresholding using a small scale example having two subnetworks with each subnetwork having a provincial hub (blue) and linked by a connector hub (orange), shown in (a). In (b), warmer color indicates higher connectivity and black dots indicate the ground truth adjacency matrix. We denote $\bar{\mathbf{C}}$ as global thresholded, and $\hat{\mathbf{C}}$ as local thresholded connectivity matrix. At a graph density of 0.25, the Global Thresholding (GT) generated isolated node 2 in (c), while our Local Thresholding (LT) preserved two edges linked to node 2 in (d).	86
Figure 3.11	Will90fROI atlas [23] with 90 region of interest (ROI)s being assigned to 14 well-established brain systems.	87

Figure 3.12	Subnetwork extraction accuracy using synthetic dataset using different thresholding schemes. The bar chart indicates the average DSC over 100 synthetic dataset across a graph density range of [0.005, 0.5] at an interval of 0.01.	88
Figure 3.13	Circular plots of the whole brain connections based on the 14 subnetwork structure using different thresholding schemes. Exemplar results shown here are thresholded at the graph density of 0.02. The thickness of the connections indicate the strength of the connections.	89
Figure 4.1	Schematic illustrating our method using small scale example having two subnetworks with each subnetwork having a provincial hub (blue) and linked by a connector hub (orange), shown in (a). In (b), warmer color indicates higher connectivity and black dots indicate the ground truth adjacency matrix. We denote \bar{C} as global thresholded, and \hat{C} as local thresholded connectivity matrix. At a graph density of 0.25, the GT generated isolated node 2 in (c), while our LT preserved two edges linked to node 2 in (d). Refining the graph (c) and (d) suppressed the between-network edges (edges between nodes 6 and 7 & nodes 6 and 9) to be the lowest connectivity in (e) and (f).	96
Figure 4.2	Subnetwork extraction on synthetic data at graph densities from 0.005 to 0.5 at interval of 0.01. Our proposed strategy attained the highest DSC overall.	97
Figure 4.3	Subnetwork extraction on real data at graph densities from 0.05 to 0.5 at interval of 0.05. Our proposed strategy attained the highest DSC overall.	98

Figure 4.4	Subnetwork visualization. 11 subnetworks were extracted from graphs with a density of 0.2. (a) Well-established brain systems [23]. (b) Two subnetwork formed by isolated nodes and false inclusion of premotor-related regions into auditory system was observed using global thresholding. (c) Local thresholding failed to detect one region of known visual systems and falsely detected four unrelated regions into dorsal default mode system. (d) Our strategy LTMR correctly detect most of the subnetworks found in [23].	99
Figure 4.5	Augmented graph model for our augmented RW with prior model. The use of prior is equivalent to using M labeled “floating” augmented nodes (dash-line circles in black) that correspond to each label and are connected to each blue node with black dash lines.	103
Figure 4.6	Illustration of robustness of using multiple seeds by mitigating the problems caused by noisy connections between subnetworks. Noisy connections between subnetworks could affect the power of single seed by creating strong “false” walks (thick purple line connecting the blue seed to a node which should be in the red subnetwork). By using multi-seeds (the smaller floating dots around the big dot in the center), we can preserve several reliable connections (the dash lines connecting those smaller dots to the nodes) to battle against possible false positive connections.	105
Figure 5.1	Schematic illustration of multi-modal Random Walker (mmRW), where the graph weights in the previous iteration $t - 1$ are defined by FC edges in blue and the seeds are found using the provincial hubs derived from AC (shown as red dash circles), then the weights in the current iteration t are defined by AC shown in red edges and the seeds are found using the provincial hubs derived from FC (shown as blue dash circles). . .	114

Figure 5.2	Subnetwork extraction accuracy on synthetic data. Our mmRW approach achieved significantly higher DSC than unimodal and existing multimodal methods.	117
Figure 5.3	Subnetwork extraction overlap to well-established brain system [23] on real data from HCP. Our mmRW approach achieved significantly higher DSC than contrasted methods.	118
Figure 5.4	Inter-subject reproducibility on real data from HCP. Multimodal RW approach achieved comparable DSC to unimodal RW approach, but significantly higher DSC than existing multimodal methods.	119
Figure 5.5	Subnetwork extraction results on an exemplar brain graph. (a-b) Probabilities of a node being assigned to a given subnetwork (color-coded in 14 subnetworks), the larger size of a node indicates a higher probability. (c-e) Exemplar subnetworks identified by mmRW.	120
Figure 5.6	Hypergraph and its corresponding simple graph and incidence matrix. Left: an hyperedge set $E = \{e_1, e_2, e_3, e_4\}$ and a node set $V = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$. Middle: the corresponding simple graph. Right: the incidence matrix \mathbf{H} of the hypergraph on the left, with the entry (v_i, e_j) being set to 1 if v_i is in e_j , and 0 otherwise.	125
Figure 5.7	Subject-wise level modularity Q values using Method (1)-(6). For method (5) and (6), parameter γ and ω were selected at the highest inter-subject reproducibility.	135
Figure 5.8	Subject-wise level inter-subject reproducibility of subnetwork extraction using Method (1)-(6). For method (5) and (6), parameter γ and ω were selected at the highest inter-subject reproducibility.	136
Figure 5.9	Visualization of subnetworks extraction using methods (1)-(6). The brain is visualized in the axial view. The mass center of each ROI is plotted in the Montreal Neurological Institute (MNI) space and color-coded by the membership of seven subnetworks.	138

Figure 5.10	The schematic illustration of multisource clique based overlapping subnetwork extraction approach.	145
Figure 5.11	Group-level Subnetwork Extraction reproducibility based on data from two different sessions. Multisource Clique-based Subnetwork Extraction (MCSE) outperforms all other contrasted approaches.	153
Figure 5.12	Subject-wise level inter-subject reproducibility of subnetwork extraction. Our proposed MCSE approach outperforms existing state-of-the-art overlapping community detection methods.	154
Figure 5.13	Visualization of Task activation and overlapping subnetworks extracted from our proposed approach and contrasted three other methods. The brain is visualized in the axial view. Our proposed MCSE approach outperforms existing state-of-the-art overlapping community detection methods by detecting well-known hubs which reside within subnetwork overlaps.	155
Figure 5.14	Overlapping confidence of interacting nodes in blue versus individual nodes in red derived by MCSE with C^{t-f} . The probability of a node being assigned into subnetworks was derived by the RW based approach [6].	159

List of Acronyms

AAL Automated Anatomical Labeling

AC Anatomical Connectivity

ACC Anterior Cingulate Cortex

AD Alzheimer's disease

BOLD Blood Oxygenated Level Dependent

BP Boundary Point

CBP connectivity based parcellation

Cg Cingulate Cortex

CIS Connected Iterative Scan

CNN Convolutional Neural Networks

COREG CO-training with REGularization

CPM Clique Percolation Method

CSA Constant Solid Angle

CSORD Coupled Stable Overlapping Replicator Dynamics

DAE Deep Auto-Encoder

DALYs Disability-Adjusted Life Years

DBN Deep Belief Network

DC physical distance

DMN Default Mode Network

dMRI Diffusion-weighted Magnetic Resonance Imaging

DNNs Deep Neural Networks

DOF degree of freedom

DSC Dice Similarity Coefficient

DTI Diffusion Tensor Imaging

EC Effective Connectivity

ECN Executive Control Network

EEG Electroencephalography

EPI Echo-Planar Imaging

FA Fractional Anisotropy

FC Functional Connectivity

fMRI Functional Magnetic Resonance Imaging

GBD Global Burden of Disease

GLM General Linear Model

GMM Gaussian Mixture Model

GROUSE Grassmannian Rank-One Update Subspace Estimation

GT Global Thresholding

HARDI High Angular Resolution Diffusion Imaging

HCP Human Connectome Project

HO Harvard-Oxford

ICA Independent Component Analysis

IP Interior Point

IPL Inferior Parietal Lobule

IQ Intelligence Quotient

LECN Left Executive Control Network

LMaFit Low-Rank Matrix Fitting

LT Local Thresholding

MC Multimodal Connectivity

MCNF Matrix Completion with Nonnegative Factorization

MCSE Multisource Clique-based Subnetwork Extraction

MEG Magnetoencephalography

MITK Medical Imaging Interaction Toolkit

mmRW multi-modal Random Walker

MNI Montreal Neurological Institute

MRI Magnetic Resonance Imaging

MST-KNN minimal spanning tree and k-nearest neighbors

MVSC MultiView Spectral Clustering

Ncuts Normalized cuts

NMD Neighborhood-information-embedded Multiple Density

NMF Non-negative Matrix Factorization

NMI Normalized Mutual Information

NP Non-deterministic Polynomial-time

NRMSE normalized root-mean-squared-error

ODF Orientation Distribution Function

OSLOM Order Statistics Local Optimization Method

PCA Principal Component Analysis

PCC Posterior Cingulate Cortex

PD Parkinson's disease

PDD Principal Diffusion Direction

PDF Probability Density Function

PET Positron Emission Tomography

RBM Restricted Boltzmann Machines

RD Replicator Dynamics

RM random parcellation

ROI region of interest

RQ Research Questions

rs-fcMRI Resting State Functional Connectivity based on MRI

RW Random Walker

sMRI Structural Magnetic Resonance Imaging

SNR signal-to-noise ratio

SORD Stable Overlapping Replicator Dynamics

SVR Support Vector Regression

t-fcMRI Task Functional Connectivity based on MRI

TR repetition time

WHO World Health Organization

3D three-dimensional

Acknowledgments

I would like to thank my advisor, Dr. Rafeef Abugharbieh, for supervising me throughout my doctoral studies. Her guidance and support have encouraged me to make it through every tough moment. I owe particular thanks to Dr. Bernard Ng for his guidance to research. Furthermore, I would like to thank my friends inside and outside of BiSICL for their support. Last but not least, I would like to thank my parents and my boyfriend, Logan, for always being there for me.

Chapter 1

Introduction

If the human brain were so simple that we could understand it, we would be so simple that we couldn't. — Emerson W. Pugh

The brain is the command center for the human nervous system and controls most of the activities of the body. It is one of the human body's most complex and magnificent organs. It contains hundreds of billions of neurons (grey matter), which are connected by trillions of connections (synapses), and billions of nerve fibers (white matter). The brain's connectivity pattern is represented by anatomical links, such as synapses, nerve fiber pathways, or functional links which reveal the information flow between different regions. Neural activities are constrained by connectivity and the interrelations between anatomical and functional aspects. Brain connectivity is thus crucial to elucidating how neural networks behave to enable complex cognitive processes.

1.1 Motivation

According to the latest Global Burden of Disease (GBD) study, published by the World Health Organization (WHO) in October, 2016, neurological and mental disorders are leading causes of suffering and disability, affecting 2.26 billion people worldwide [10]. To quantify the effects of different diseases, the GBD study designed a metric termed Disability-Adjusted Life Years (DALYs) to measure the number of years of life lost as a result of both premature death and disability. Based

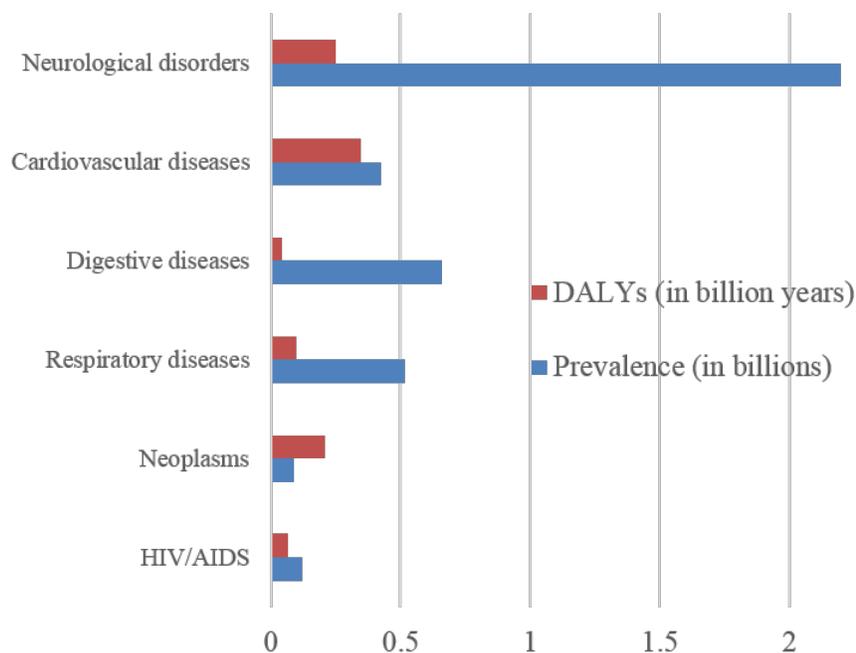


Figure 1.1: Prevalence and total DALYs for selected diseases and neurological disorders. Data are adopted from [9, 10].

on the latest GBD data for 2015, neurological and mental disorders cause 250 million DALYs, and this demonstrated a 36% increase since 1990 [9]. A comparison of the prevalence and DALYs for selected common diseases is given in Figure 1.1. In Canada, based on Mapping Connections report released by Neurological Health Charities Canada in September, 2014, 3.6 million people (more than 10% of the total population) are affected by neurological disorders [24]. By 2031, the number of Canadians with neurological conditions will increase, and the number of people with Alzheimer’s disease (AD) and other dementias, Parkinson’s disease (PD) and traumatic brain injury is expected to double [24]. This report, through microsimulation results, predicts that Canadians born during the current decade (2010-2020) who will be affected by neurological disorders will lose between 14 and 41 years of fully healthy lives [24]. Having a neurological condition affects one’s general and mental health, causes impairments in function that impact on the quality of

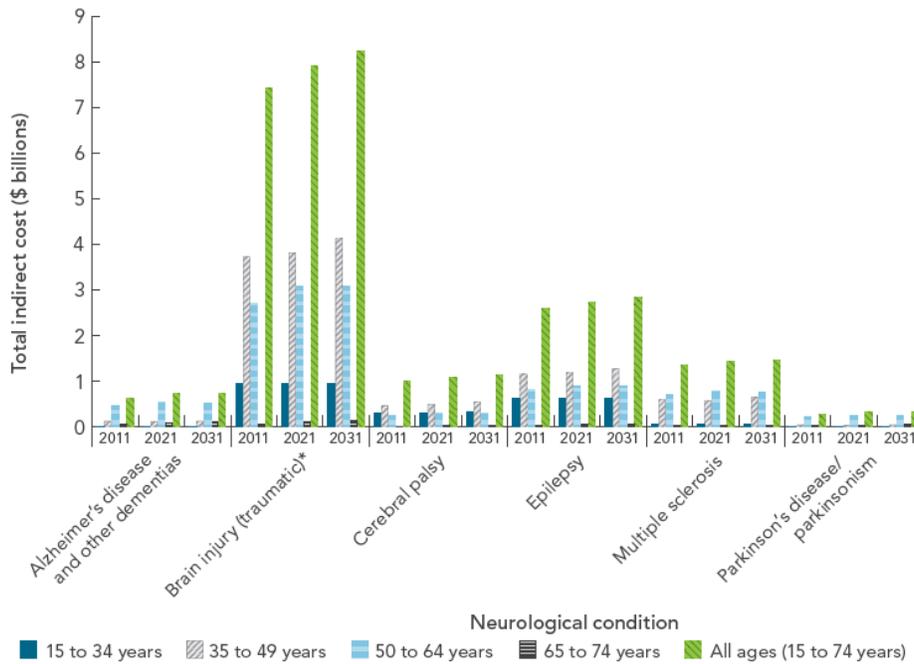


Figure 1.2: Projected indirect economic costs due to working-age death and disability, by select neurological condition and age group, Canada, 2011, 2021, and 2031

life, and induces a tremendous economic cost. The mapping connections study predicted that these costs are expected to progressively increase as the population life span extends (Figure 1.2). Recent studies indicate that many neurological disorders can be described as dysconnectivity syndromes, associated with changes in the structural and functional brain networks prior to the development of clinical symptoms [25]. Our enhanced understanding of brain network organization can help explain network changes in neurological diseases. As a result, there is a genuinely urgent need to comprehend the brain's connectivity and the interaction of its regions, and to reveal these relationships to diseases of current significant fundamental importance.

1.2 Thesis Objectives

The overarching research objective of this thesis is to develop methods for a multimodal analysis of human brain imaging data to better understand the brain’s structure, function and connectivity, i.e., brain connectomics. By studying the relationship between brain structure and function, we propose multimodal fusion approaches for human brain connectivity analyses based on graph theory. Our rationales are that, integrating information from various sources would be beneficial by combining their respective strengths; additionally graphical metrics should play an important role in brain network analysis.

In this thesis, we focus on studying the healthy brains to build a baseline of how regular brain network works. Our aim is to discover the intrinsic features of the brain without being biased by discriminative patterns in certain diseases. The improved understanding of normal brain-network organization could make it possible to explain network changes in a wider range of neurological and psychiatric diseases. It is important to study healthy brains and uncover network changes prior to the development of clinical symptoms, which enables us to predict the disease onset and perform interventions for at risk subjects.

The resulting methods we develop could provide important new insights into cognitive and clinical neuroscience, where the findings may be applied to neurological disease research in revealing biomarkers for early diagnoses, prognoses, neurosurgical planning and therapeutic interventions. We believe that our research will enable further exploration of linking brain network features to the stages of disease severity, our thoughts on promising directions will be summarized in the Conclusion (Chapter 6).

1.3 Imaging for Brain Connectomics

1.3.1 Definition of Brain Connectomics

“Connectomics” is the study of brain connectivity, a term which was introduced in 2005 to refer to the complete mapping of structural connections among neurons and brain areas [26]. The term has recently come to include the mapping of both anatomical and functional connectivity [14].

1.3.2 Imaging Resolution in Brain Connectomics

The connectome can be analyzed by different scales associated with levels of spatial resolution in brain imaging. This resolution ranges from the microscale at the micrometer resolution level, which maps individual neurons and their synaptic connections; mesoscale at the hundreds of micrometers resolution level, which encompasses neuronal populations, formed by local circuits (e.g. cortical columns) that link hundreds of individual neurons; and the macroscale at the millimeter resolution division, which parcellates large brain systems into anatomically or functionally distinct brain areas. Currently, macroscale connectomics is the most commonly used scale due to its lower computational and analytical demands, which make comprehensive mapping for noninvasive and *in vivo* imaging of the whole-brain connectome feasible [14].

1.3.3 Functional Imaging for Macroscale Connectomics

The functional perspective of macroscale connectome is coded for by the temporal dependence between the activities from diverse brain areas.

1.3.3.1 Modalities for Functional Imaging

Different modalities are available for macroscale functional imaging. Electroencephalography (EEG) and Magnetoencephalography (MEG) are well suited to studying the temporal dynamics of direct neuronal activities within milliseconds of resolution. However, it is difficult to uniquely identify the location of underlying bioelectric/biomagnetic activity based on the scalp topology of measured signals, which is termed as the ill-posed inverse problem [27]. Positron Emission Tomography (PET) is an imaging technique that detects molecular biological details using radiolabelled molecular probes, based on the assumption that areas of high radioactivity are associated with brain activity [28]. Functional Magnetic Resonance Imaging (fMRI) is an *in vivo* imaging technique measuring changes in brain function via Blood Oxygenated Level Dependent (BOLD) contrast [29]. fMRI has high spatial (millimeters) and intermediate temporal (hundreds of milliseconds or seconds) resolution. Spatial-temporal resolution plots for commonly used functional neuroimaging methods have been illustrated in Figure 1.3. In fMRI read-

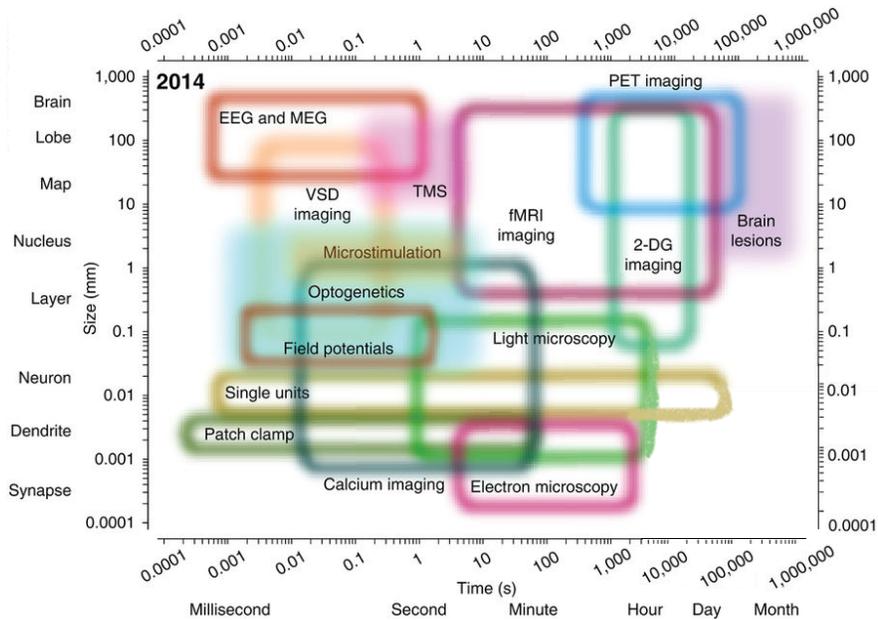


Figure 1.3: Spatial-temporal resolution plots for commonly used functional neuroimaging methods in the year of 2014. VSD, voltage-sensitive dye; TMS, transcranial magnetic stimulation; 2-DG, 2-deoxyglucose. Image courtesy of [11].

ings, the hemodynamic response, i.e., variations in blood oxygenation during neural activity, causes magnetization changes that can be detected in a Magnetic Resonance Imaging (MRI) scanner [29]. Specifically, neuronal firing consumes oxygen, which increases local magnetic inhomogeneity, and, hence, decreases BOLD signals near the brain areas involved. When an inflow of oxygenated blood rushes into these brain areas in response to the oxygen demand, the relative decrease in deoxygenated blood reduces local magnetic inhomogeneity, which increases the BOLD signals [30]. The BOLD-fMRI, which employs high localization power, is an indirect measure of brain activity through oxygenation and blood flow [31].

Among the modalities used for macroconnectomics, we have chosen the BOLD-fMRI due to its widespread availability, superior spatial resolution, good temporal resolution, and safe and noninvasive sampling fashion [14].

1.3.3.2 fMRI Experimental Design

There are broadly two types of fMRI experiments for functional connectivity studies. For the Task Functional Connectivity based on MRI (t-fcMRI), subjects are asked to perform certain cognitive tasks which help us identify and characterize functionally distinct regions in the human brain. The data for Resting State Functional Connectivity based on MRI (rs-fcMRI), are obtained when subjects are scanned without performing any explicit task requirements. Hence, the latter reflects spontaneous neural activities in the absence of bias towards any particular cognitive or motor task demands [32].

In this thesis, these two types are considered as two different sources, with the rs-fcMRI being used for standard connectivity analyses, and the t-fcMRI for the multisource fusion of rest and task data using both the connectivity derived from the task data and the activation magnitude or location of important brain regions.

1.3.4 Structural Imaging for Macroscale Connectomics

The structural perspective of the macroscale connectome is an organizational description of the network of elements and connections forming the human brain. This structural or physical pattern of connections is relatively fixed at shorter time scales. There have been attempts to map the large-scale structural architecture of the human cerebral cortex, such as by exploiting correlations in cortical volume or thickness using Structural Magnetic Resonance Imaging (sMRI) [33]. Cortical thickness correlations have been postulated as being indicators for the presence of physical structural connections. However, this approach provides highly indirect information about connection patterns. It is worth mentioning that, currently, only invasive tracing studies are able to illustrate direct axonal connections. Diffusion-weighted Magnetic Resonance Imaging (dMRI) has been primarily used to uncover relatively direct *in vivo* markers of fiber tracts with whole-brain coverage.

The dMRI estimates white matter integrity and connectivity by examining the translational displacement of water molecules [34]. Water molecules are in constant thermal motion, constrained by physical boundaries. In fibrous tissues, such as axons in white matter, water diffuses more rapidly in the direction aligned with the structure and more slowly in the perpendicular direction. Measurements of

this anisotropic diffusion thus reveal micro-structural properties of the underlying tissue. In practice, images sensitized to different diffusion directions are acquired, followed by a fiber reconstruction step to estimate the voxel-wise three-dimensional (3D) diffusion Probability Density Function (PDF) [35]. Based on the PDF, tractography is a methodology capable of generating a continuous, smooth representation of the white matter fiber trajectories.

1.4 State-of-the-art Techniques in Brain Connectivity Analysis

1.4.1 Brain Network Analysis for Connectivity Studies

The human brain can be regarded as being a network where units, or nodes, represent different specialized regions, and edges represent communication pathways. Brain network analysis methods for connectivity studies can be classified into two types: model-dependent methods for seed-based network analysis, and model-free methods for whole brain network analysis [36].

1.4.1.1 Model-dependent Methods for Seed-based Network Analysis

Model-dependent analysis is centered on a seed region, which may be identified either as a lesion, a region of interest (ROI) involved in a specific pathway, or a function of a brain disorder [37]; it can also be selected from a traditional task-dependent activation map acquired from meta-analysis results from previous fMRI experiments [36]. In the seed-based analysis, signals from only the voxels within the seed are used to compute the functional correlations between time series of the seed region and other voxels using fMRI data; they can also be utilized to track the fiber tracts passing the seed region using dMRI data [37]. This provides a precise and detailed description of regions of specific connectivity in particular areas of interest. For example, Marwan *et al.* extracted six seed regions to study how chronic pain disrupts the Default Mode Network (DMN) [38]. This is one of the well-known networks which is more active at rest than during task performance, suggesting the existence of a resting state in which the brain remains active in an organized manner. The seed-based methods have the advantage of producing

straightforward results. However, these types of methods only examine limited connectivity information concerning selected ROIs, and the results might depend on the selected seeds and are thus greatly dependent upon the experience and prior knowledge of the researcher [37].

1.4.1.2 Model-free Methods for Whole Brain Network Analysis

Model-free methods aim at identifying whole brain networks in a data-driven manner. The related methods are used to discover general patterns of connectivity across brain regions without defining a *prior* ROIs [36]. Based on connectivity information, methods such as matrix factorization, clustering, and graph-theoretic approaches have been applied to whole brain network analysis. For example, matrix factorization techniques such as Principal Component Analysis (PCA) [39], Independent Component Analysis (ICA) [40] and Non-negative Matrix Factorization (NMF) [41] have been applied to whole brain network analysis. The main drawback of such methods is the lack of interpretability of the resulting components (independent or principle components), since they contain a more complex representation of the data [36]. Clustering methods such as k-means [42], hierarchical clustering [43] and Laplacian clustering [44] target at maximizing the similarities between the data points within connected sub-clusters, i.e., subnetworks [36]. In graph-theoretic approaches, a graph is an intuitive and straightforward representation of a brain which is organized by grey matter regions (nodes) connected by white matter tracts (edges). Graph theory provides a theoretical framework for examination of the topology of networks and reveals both the local and global organization of brain networks [18, 45].

In order to explore the connectivity of the whole brain network (connectomics), we chose to focus on the graph theory, which provides a compelling framework for the analysis of large-scale brain network architecture. The graph theory framework, which has been widely used in cognitive and clinical neuroscience [46], enables comprehensive studies on brain connectivity from three levels: candidate circuit analysis (similar to seed based approaches), connectome-wide analysis (similar to matrix factorization applications) and topological analysis (unique to graph theory techniques) [47].

1.4.2 Graph Representation in Brain Connectomics

In the graphical representation of the brain connectome data, the brain network can be defined as a graph $G = (V; E)$ with V being the set of nodes reflecting the brain regions, and E corresponding to the degree of interactions between the brain units modeled by edge weights.

1.4.2.1 Existing Studies of Graph Use for Brain Connectomics

“Network neuroscience” has been proposed to refer to research where neuroscience components are studied mathematically in the form of networks or graphs [46]. Graphs have been used in early neuroscience research on non-human brain networks, such as the seminal work examining the organization of large-scale connectivity networks in the *Caenorhabditis elegans*, which possesses 302 neurons and 6,393 connections [48]. “CoCoMac” comprises approximately 40,000 experimental findings on anatomical connections in the macaque brain, as derived from neuroanatomical tract tracing studies [12], Figure 1.4. Most of these studies were conducted by graph-theoretical analyses of the network structures of primates.

Recent studies also conceptualize the human brain as a graph organized across different spatial and temporal scales, i.e., a distributed complex system using graph theory whose integrated function underlies human behaviour and cognition [16, 45]. There has been a great and rapidly growing number of studies on both healthy and disease brains using graph theory during the past decade, detailed in review papers ([49–51] for healthy brain, and [47, 51, 52] for disease brain).

These studies have indicated that healthy brain systems have features of complex networks, such as small worldness, modularity and rich club of hubs [16]. Brain graph organization in neurological disorders almost always reflects a deviation from the normal pattern, which is characterized by high local connectedness (indicated by high clustering coefficient), high integration (short path length), hierarchical modularity, and hub nodes that are interconnected in a rich club [52].

A significant number of evidence indicates that graph theory enables us to discover the mechanisms of normal brain organization and functions, uncover developmental mechanisms leading to abnormal brain network organization, and track the progression of disease in neurological disorders [47].

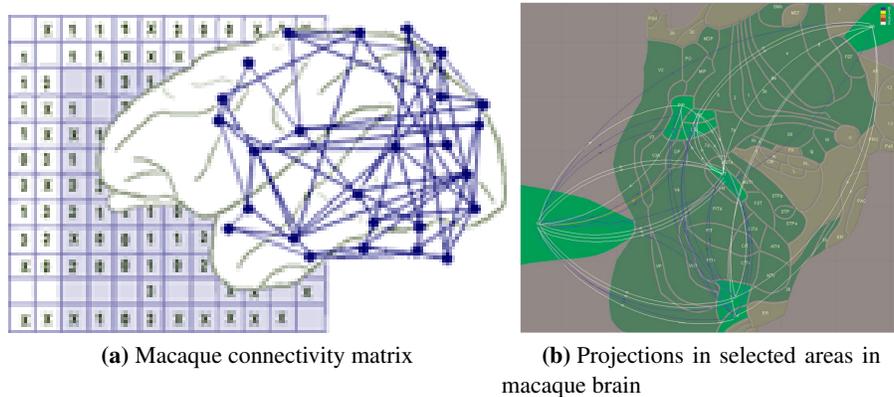


Figure 1.4: Connectivity matrix based on graph theory created by Rolf Kötter and projections derived from tract tracing, one of the first attempts to obtain a large-scale anatomical connectome of the mammalian brain. Image courtesy of [12].

1.4.2.2 Definition of Nodes: Parcellation-based Brain Connectome

In graph analysis frameworks, valid node definition, i.e., parcellation, is critical for the accurate mapping of inter-regional connectivity. Inconsistent or imprecise node identifications will have a huge impact on subsequent analyses [53]. Accurate parcellation enables an efficient comparison of results across studies, and it acts as a foundation for illuminating the structural and functional organization of the brain; and as a means to reduce data complexity while increasing statistical sensitivity and power [54]. Ideal brain nodes should represent homogeneous brain regions, retain functional heterogeneity with the other nodes, and ensure spatial contiguity [16].

Whole-brain parcellation still remains challenging in the neuroimaging field [16] due to multiply intertwined problems. First, what properties define a brain unit are unclear. For MRI data, a voxel is often considered a brain unit [55], but this definition solely depends on the scan protocol, e.g., a 1 vs. 2 mm. acquisition would result in “neuron patches” of different sizes being defined as brain units. Another alternative is random parcellation, which typically produces roughly equisized parcels, but does not guarantee within-parcel homogeneity [56]. This disadvantage can be partially alleviated by the generation of a very large number of parcels; how-

ever, this in turn significantly increases computational complexity in subsequent analysis steps. A common and likely more neuroanatomically-meaningful strategy is to group voxels based on certain neurobiological properties, such as gyral-based architecture [57] (well-known examples include the Harvard-Oxford (HO) atlas), cyto-architecture [58], myelo-architecture [59], chemo-architecture [60], cortical folding [61], shared brain function [62], regional homogeneity [63], meta-analytic activation modeling [64] and connectivity patterns, i.e., connectivity based parcellation (CBP) [65–67]. With today’s wide availability of MRI scanners, most strategies are now designed around using only MRI-based features for parcellation [56, 65, 66]. In particular, the majority of recent works focus on CBP, which does not require generating a very large number of parcels, since within-parcel homogeneity is inherently enforced, facilitating the investigation of brain networks at computationally feasible resolutions. Hence, we will focus on CBP approaches to meet the challenging demands for parcellation.

A related problem is the choice of the number of parcels. Often, this number is arbitrarily set [53]. More principled strategies include using criteria, such as predictability on leftout data [68] or stability over data splits [69]. Based on architectonic [70] and functional [71] information, five hundred parcels appear as probable. Others argue that this choice should be application-specific [72]. For example, estimating a 500×500 connectivity matrix given the presence of 100 time samples would result in a high degree of estimation errors [73].

1.4.2.3 Definition of Edges: Estimating Connections Between Brain Regions

The edge which constitutes a continuous measurement of connectivity between brain regions is usually defined by three broad classes of brain connectivity: anatomical, functional and effective [46]. In the sense of Anatomical Connectivity (AC), edges can be connection probability between two regions of tractography-derived datasets [74]. On the other hand, Functional Connectivity (FC) is most commonly computed using the Pearson’s correlation coefficient between regional activity time courses. Lastly, Effective Connectivity (EC) infers the causal interactions between two brain regions by estimating spectral coherence or Granger causality measures [75]. We focus on AC and FC in this thesis for the reason that current tempo-

ral resolutions of available fMRI data might not suffice to perform accurate EC estimations.

1.4.2.3.1 Estimating Anatomical Connectivity The underlying principle of estimating anatomical connectivity from tractography-derived data is to rely on water diffusion as an indirect probe of axon geometry [13]. By tracking the motion of water, one can map the orientation(s) of fibers passing through each voxel of white matter [14]. Since each voxel contains thousands of axonal fibers, one is first to infer a probability function for each voxel as the fiber Orientation Distribution Function (ODF), which captures the different fiber orientations present [14]. A simple model that approximates the fiber ODF is Diffusion Tensor Imaging (DTI), which uses a 3×3 positive semi-definite matrix to provide an ellipsoid representation (the major eigenvector) of the water-diffusion profile for a given voxel [34]. Based on the assumption that the direction of least hindered diffusion, or Principal Diffusion Direction (PDD), is aligned with that of the axons; tractography algorithms based on DTI, such as “streamline” tractography, have been proposed to estimate tract trajectories by connecting the ellipsoids “pointing” at each other from voxels [76]. 3D trajectories, referred to as streamlines, are used to trace putative white-matter paths. PDD provides a good estimate of fiber orientations when the axons are aligned in parallel within a voxel. In reality, fibers are known to fan/merge, cross, kiss and bend within a single voxel (Figure 1.5), which leads to a heterogeneity not accounted for by a simple ellipsoidal ODF [13]. More elaborate models for approximating fiber ODF have been used to extend the concept of a single PDD to mixtures of PDDs to include features like fanning and polarity (Figure 1.5), based on High Angular Resolution Diffusion Imaging (HARDI) techniques [35]. This process is defined as qBall imaging reconstruction since the HARDI data are usually acquired on one or multiple spherical shells in q-space [35]. These more complex ODFs can be estimated by a linear radial projection of the PDF [77] or by the projection in a cone of Constant Solid Angle (CSA) [78]. It has been demonstrated that using the projection of CSA provides a sharper representation of the diffusion process, resulting in a clearer reconstruction of certain fiber bundles [78]. Using more sensitive ODF models does not guarantee more accurate tractography results. For these more complex ODF models, streamlining

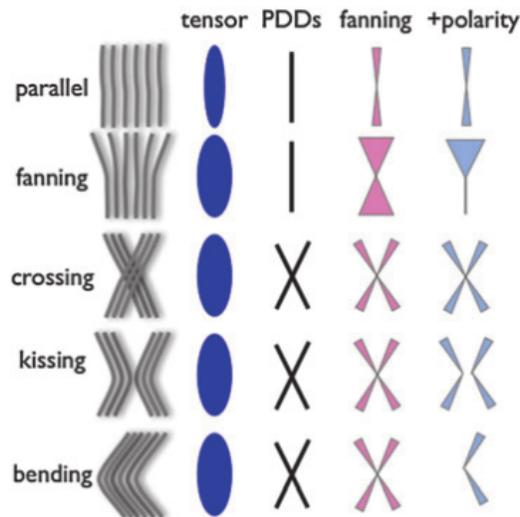


Figure 1.5: Cartoon illustrations of ambiguities in mapping diffusion to axon geometry. Different axon geometries can lead to a similarly oriented tensor. The tensor’s principal direction is the same for all cases, but modeling multiple PDDs helps distinguish a few of the cases. Modeling fiber fanning separates the top two cases. Further modeling the polarity of a fanning can help separate all cases. Image courtesy of [13].

follows the same principle, but with multiple peak orientations available at each voxel rather than a single PDD, which enables streamlines with different orientations to pass through the same voxel. However, these streamline algorithms are also prone to errors since they require the estimation of the number of bundles in a voxel to determine the number of ODF peaks. Measurement noise usually induces spurious variations in the generated streamlines [14]. Additionally, use of a fixed step size in the generation of streamlines (despite local variations in anatomy where all the configurations exit within a voxel) aggregates measurement error when traditional tractography approaches are deterministic. Probabilistic tractography algorithms model these errors or noise by estimating the uncertainty of local fiber orientations in the entire process [14]. Uncertainty in voxel-wise fiber orientation can be quantified regarding the locations of streamlines [79], Figure 1.6. Nevertheless, these algorithms are mostly very computationally expensive, which limits their use for whole-brain connectivity. An alternative method that bypasses

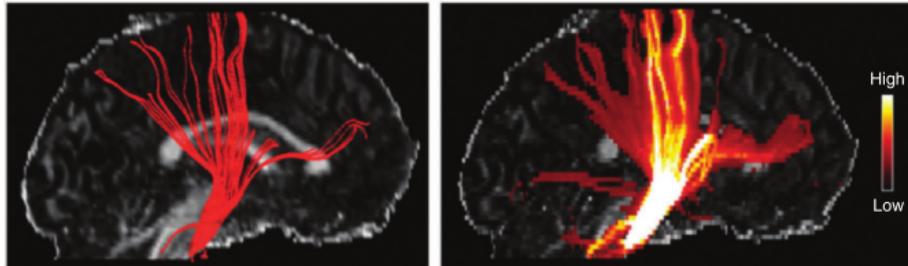


Figure 1.6: Deterministic and probabilistic tractography. Pyramidal tract streamlines based on deterministic (left) and probabilistic (right) approaches (results superimposed on FA image). Color bar indicates confidence about the presence of the tract. Image courtesy of [14].

the potential source of error of deterministic tractography at a reasonable computation speed is global tractography [80], where the decisions are made on a global level, aggregating diffusion data across voxels to infer subvoxel features. Global tractography jointly considers all fiber trajectories in determining the most plausible fiber configuration [81]. Streamline tractography tends to terminate tracking at locations such as crossing fiber, whereas global tractography exploits the geometry of the adjacent fibers by examining the whole-brain fiber configuration simultaneously to compensate for ambiguous local information. This strategy alleviates error propagation along tracts and reveals a greater number of known tracts missed with conventional approaches, such as bilateral tracts that connect the two hemispheres [82]. Collectively, more accurate models and better signal descriptions are necessary to overcome the limitations of the state-of-the-art fiber reconstruction and tractography methods [83].

In order to quantify the connectivity strength for AC estimation based on fiber tracts reconstructed from tractography results, metrics such as uncertainty, fiber counts, fiber density, fiber length, and FA have been used [84]. We note that the uncertainty of streamline trajectories is not easily equal to connectivity strength. The approximation that stronger connections are expected to have lower uncertainty in their fiber trajectories can break in cases where locally nondominant pathways have greater degrees of uncertainty [14]. Fiber count, arguably the most common AC metric, measures the number of fiber tract streamlines connecting pairs of brain

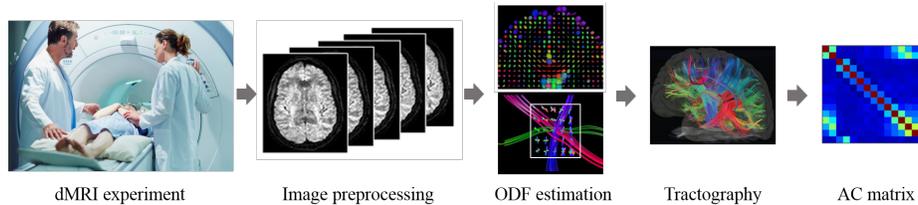


Figure 1.7: Typical workflow for AC estimation based on a dMRI experiment. Diffusion weighted images are acquired and preprocessed to reconstruct the diffusion ODFs. Fiber tracts are then generated using tractography. AC matrix can indicate the fiber counts between each brain pair. MR scanner image: A patient is loaded into an MRI machine, by Morsa Images/Getty Images, <https://www.verywell.com/diffusion-weighted-mri-3146133>, retrieved on August 30, 2017.

regions. Fiber density (fiber count / the number of voxels in brain regions), is a variant of fiber count conducted by normalizing the fiber count by the total volume of the brain region pairs interconnected in numbers of voxels. Mean fiber length (i.e. the average length in mm for all streamlines belonging to a fiber tract) is sometimes used to correct the bias that longer tracts may have larger accumulated error, leading to lower fiber counts. The average FA along streamlines connecting brain regions has also been proposed as a proxy for connection strength [85].

Based on the previous work in our lab [86], we have chosen to use the CSA-ODF estimation and global tractography based on HARDI data. We use the normalized fiber count to quantify the AC due to its intuitive interpretation of connection strength. Steps in a typical dMRI experiment for AC estimation are shown in Figure 1.7.

AC estimated from tract strength based on dMRI data particularly suffers from false negative connections, Figure 1.8. The false positives and negatives here are not this way in the common statistical sense; instead, they are misestimates of connectivity due to modeling errors [15]. Due to partial volume effects from the limited spatial resolution, problems such as ambiguous diffusion direction at crossing fibers during tractography can cause the premature termination of tracts with conventional streamline algorithms [87]. Even when this problem is resolved with better ODF modeling, the ambiguity in diffusion direction near the gray-white mat-

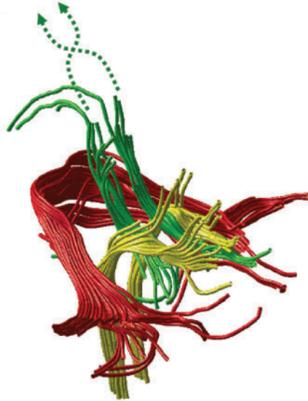


Figure 1.8: Example of false negatives in *in vivo* tractography of the cerebellum. The dotted green lines indicate the incomplete reconstruction of the superior cerebellar tracts where streamlines stop before crossing to the contralateral side as expected from known post-mortem neuroanatomy (false negatives). Image courtesy of [15].

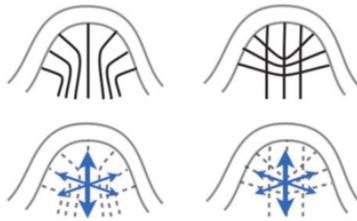


Figure 1.9: Example of ambiguities near the gray-white matter interface. Both axon configurations lead to the same diffusion profile and tracking results, but with different connectivities. Image courtesy of [13].

ter interface introduces great uncertainty concerning the fiber endpoint locations [13], Figure 1.9, which can also lead to fiber tract pre-mature termination resulting in false negatives.

1.4.2.3.2 Estimating Functional Connectivity Functional Connectivity reflects the co-activation patterns between distributed and often spatially remote brain regions during both spontaneous and task-evoked brain activity [46].

Different approaches for estimating FC measure different aspects of connec-

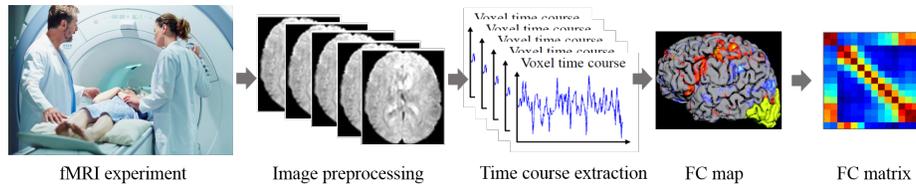


Figure 1.10: Typical workflow for FC estimation based on a fMRI experiment. Functional MRI images acquired when the subjects are at task conditions or rest. Voxel and regional time courses are extracted after image preprocessing. FC matrix can indicate covariance between each brain pair. MR scanner image: A patient is loaded into an MRI machine, by Morsa Images/Getty Images, <https://www.verywell.com/diffusion-weighted-mri-3146133>, retrieved on August 30, 2017.

tivity. The most widely used Pearson’s correlation quantifies marginal (direct and indirect) dependencies, while partial correlation measures conditional (direct) dependencies. Other metrics include spectral coherence and more complex metrics using such as phase-locking [88]. Simulation studies have demonstrated that correlation metrics perform notably better than complex metrics [88]. Real data studies on test-retest analyses have shown that correlation is more stable than coherence [89]. We have thus chosen to use Pearson’s correlation in this thesis for FC estimation. Steps in a typical fMRI experiment for FC estimation are shown in Figure 1.10.

False positive correlation is one major problem when estimating FC [90]. fMRI measurement of brain activity is indirect and error-prone, because it captures the average effect of many spikes and does not resolve cortical columns or individual neurons, resulting in structured noise and signal convolutions [91]. Hence, the inherently low signal-to-noise ratio (SNR) of the BOLD signal together with additional confounds such as scanner drift, the subject’s head motions and other physiological noise (cardiac and respiratory cycles, arterial CO_2 concentration, blood pressure/cerebral autoregulation, and vasomotion) pose major challenges to the interpretation of fMRI data, resulting in many false positive correlations in FC estimations [90].

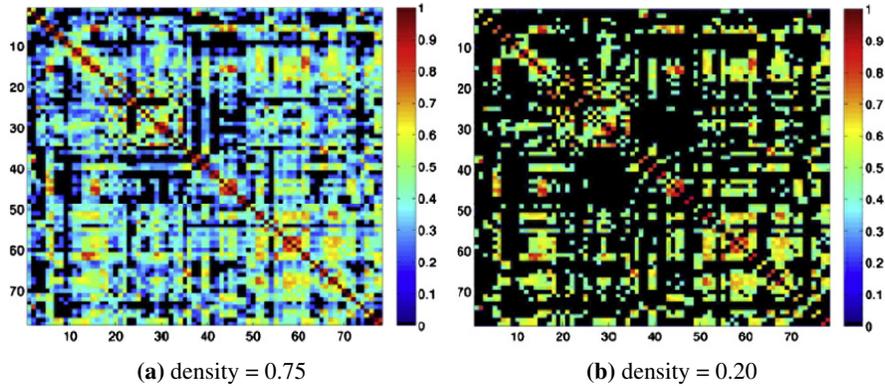


Figure 1.11: Examples of the connectivity matrices after graph density thresholding. Image courtesy of [16].

1.4.2.3.3 Thresholding Typically, measurements of both anatomical and functional connectivity strength are formed into a connectivity matrix, Figure 1.7 and Figure 1.10. The connectivity matrix can be thresholded and binarised to create an unweighted graph presented by a binary adjacency matrix. The choice of thresholds is non-trivial: different thresholds will create graphs of distinct connection density (or network sparsity), see the example in Figure 1.11. Hence, network properties are often analyzed over a range of reasonable thresholds. Meanwhile, methods deploying weighted graphs are emerging and increasingly being applied to brain network analysis [16]. In this scenario, a positive correlation between regional activity time courses represents integration or cooperation, while a negative correlation suggests competition or segregation [92]. Most classical graph theoretic metrics are based on only positively weighted connections due to the difficulty in interpreting negative connectivity [93].

Overall, it is important to leverage these different types of connectivity to define weighted edges, which potentially enables the full richness of the available multimodal data to be exploited for connectomics studies [94].

1.4.2.4 Network Measures: Graphical Metrics for Brain Connectomics

Given brain connectivity matrices, brain networks can be quantitatively examined for certain commonly used network measures of anatomical and functional con-

nectivity. Network measures are often represented in two ways. One is the measurement of individual network elements, i.e., nodes or edges, quantifying the connectivities linked to these elements to reveal the manner in which they are connected within the brain network. Subsequently, the measurement values of each metric comprise a distribution, which renders a global description of the network [17]. This distribution is typically represented by the mean value and shape, which can be used to estimate the types and properties of the network. The widely used measures detect various aspects of functional segregation and integration, quantify the importance of individual brain regions, and examine networks' resilience against attacks [17]. Segregation and integration are two principles that link different modes of brain connectivity [95]. Segregation refers to the presence of specialized brain regions with distinct functions forming segregated cortical areas, while integration indicates the collaboration of distributed neuronal populations enabling the emergence of coherent cognitive and behavioural states [95]. For example, the clustering coefficient is a good measurement of functional segregation, quantifying the presence of clusters or modules within a network. The average shortest path length between all pairs of nodes in the network, known as the network's characteristic path length, is one of the most commonly used measures of functional integration. On the other hand, in order to study the information flow or resilience of the brain network, we focus on hubs, i.e., highly connected nodes which reflect the importance of certain core brain regions. They either connect primarily with other nodes in the same group (provincial hubs) or with nodes that belong to different groups (connector hubs). Figure 1.12 illustrates some basic measures, and a comprehensive list of mathematical definitions of the measures can be found in [17].

1.4.2.5 Module Detection: Brain Subnetwork Extraction

Among the above measures, the modular structure (community structure) is of particular interest; it is from this structure that we can infer information about brain subnetworks. The modular structure is extracted by subdividing a network into groups of nodes with the maximal possible within-group links and minimal between-group links using community detection methods [96]. The term *mod-*

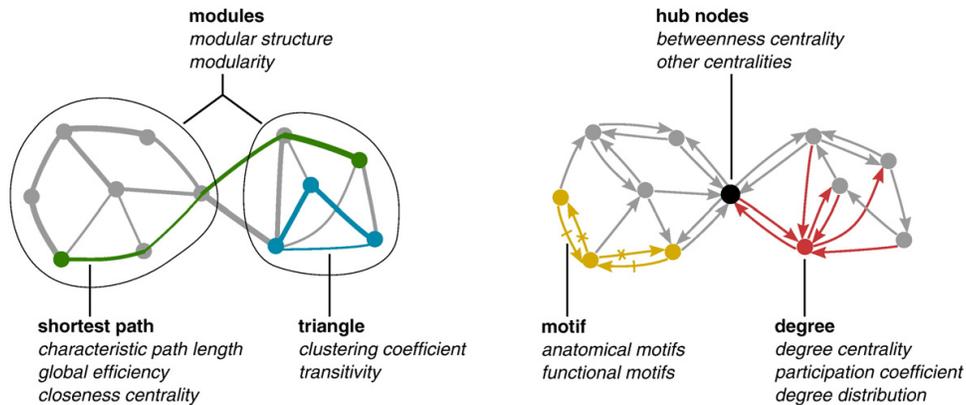


Figure 1.12: Exemplar network parameter measures. Image courtesy of [17].

ularity quantifies the degree to which the network may be subdivided into such clearly delineated and non-overlapping groups [97]. Typically, the optimal modular structure is extracted using optimization algorithms based on modularity. The notable modularity maximization approach [97] is known to be sufficiently fast for smaller networks. Another heuristic method based on modularity optimization [98] performs much faster for larger networks and is also able to detect a hierarchy of modules. There are other approaches for discovering overlapping modular network structure, where single nodes may belong in more than one specific module. For example, a new definition of modularity has been proposed based on unbiased cluster coefficients to discover overlapping subnetworks in resting state connectivity [99]. Although being able to uncover important subnetworks, the modularity measurement suffers from a resolution limit [100], which implies that the modules identified from simply maximizing their modularity might contain some smaller and more finely grained modules. This problem can be tackled by choosing different resolution parameters.

Except from modularity maximization, seed-based approach, ICA and graph partitioning approaches have been commonly used to discover subnetworks [101]. Despite its intuitiveness and simplicity, the drawback of seed-based approach is that its results are dependent on the choice of seeds. ICA based approaches are data-driven by mostly dividing fMRI observations into maximally independent spatial components [102]. At the intra-subject level, voxels that have been assigned

component weights higher than the threshold within the same spatial component are grouped into subnetworks. However, this threshold is usually set in an *ad hoc* manner, which hinders statistical interpretation across the subjects [103]. We focus on graph partitioning approaches since we use graph representations to study brain networks. Recently, graphical approaches, such as distance-based clustering techniques [64], InfoMap [18] and the Potts model [104], have been applied to the whole-brain data to extract human brain subnetworks.

It is important to note that we use subnetworks to describe the subgroups derived from a larger brain network, namely, the large-scale whole brain system. Nonetheless, we use the term “brain network” to refer to the brain system in this thesis so as to conform to the jargon used in the existing literature. Also, certain subnetworks have been denoted in the literature using jargon which sometimes includes the word “system”; however, in this thesis, they are referred to as “subnetworks”.

1.4.2.5.1 Anatomical Subnetworks In one of the earliest subnetwork studies, based on anatomical data, Hagmann *et al.*, when using spectral decomposition to maximize modularity, discovered six communities: four (two pairs of bilaterally symmetric) communities located in the frontal and temporoparietal cortex, and two positioned over the precuneus and posterior cingulate cortex [105]. One interesting finding is that these modules are spatially contiguous; this may help conserve physical wiring costs. This study also indicates a close association between the structural modules and functional domains of behaviour and cognition [106].

Two recent studies explored the multiscale human anatomical subnetworks by using a random walker moving at various timescales over the brain network [107], and by varying the resolution parameter in the multiresolution technique [108]. Both studies discovered biologically meaningful modules at different scales. Interestingly, anatomical subnetworks can be predictive of functional connectivity at some specific scales / resolutions [107].

1.4.2.5.2 Functional Subnetworks Numerous studies have shown that the human brain can be grouped into modules of functionally interconnected regions that

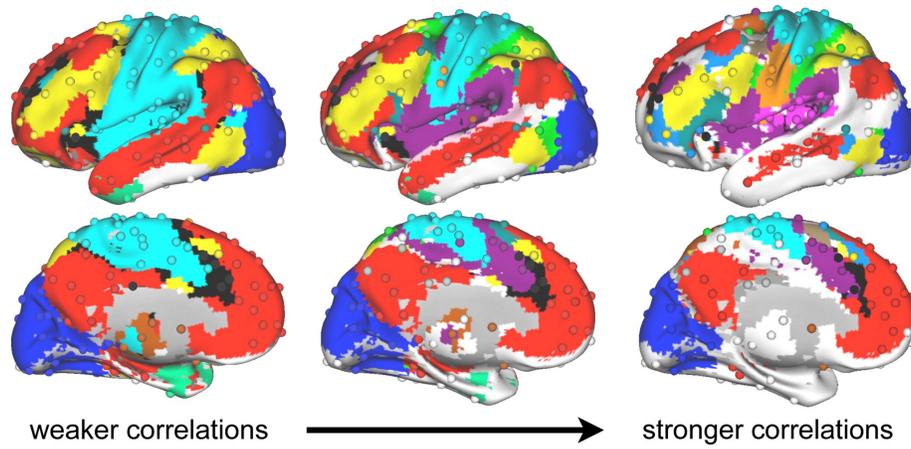


Figure 1.13: Subnetwork structure based on functional connectivity. Subgraphs from three thresholds are shown for the areal (spheres) and modified voxel-wise graphs (surfaces). Subnetworks such as default mode subnetwork, frontoparietal executive control subnetwork, and attention subnetwork comprise of spatially distributed brain regions. Image courtesy of [18].

are reproducible within subjects and consistent across subjects [109]. Power *et al.* [18] extracted 25 subnetworks from rs-fcMRI data and mapped them to cognitive or behavioural functions based on task-based activations, Figure 1.13. In this study, they have discovered some functional brain subnetworks that are not generally present in anatomical brain networks. In terms of functional connectivity, connections reflect statistical dependencies between brain regions rather than physical linkages in anatomical connectivity. Hence, functional connections do not carry direct metabolic wiring cost, which imposes a penalty on strong long distance connections [106]. Therefore, most subnetworks obtained based on functional data are spatially distributed. For instance, subnetworks such as the default mode, frontoparietal executive control, and attention subnetworks are comprised of spatially distributed brain regions [18], Figure 1.13.

Other than using rs-fcMRI, a study was able to generate a brain coactivation network based on a meta-analytic approach, using the t-fcMRI [19]. Unlike the traditional resting state FC, the foci information from task-based brain activations taken from thousands of experiments was used to build a coactivation matrix. The

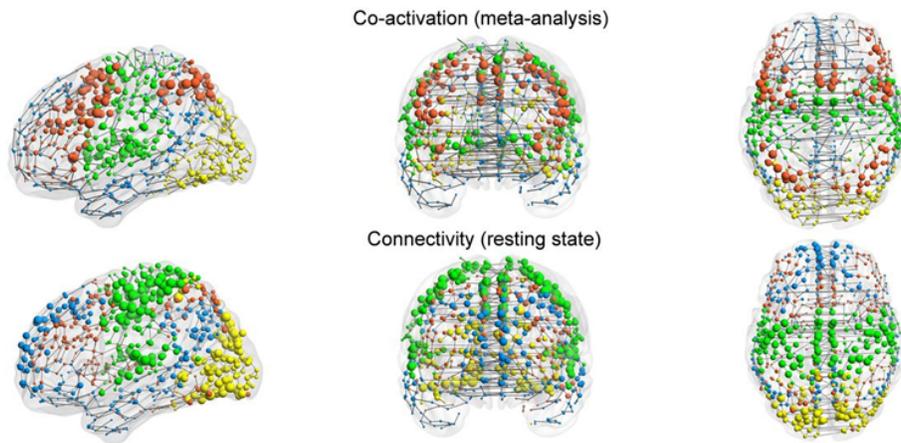


Figure 1.14: The functional coactivation network based on meta-analysis of task-related fMRI studies has similar modularity to a functional connectivity network based on resting-state fMRI data. Co-activation and connectivity networks plotted in anatomical space. The edges are defined by the minimum spanning tree for illustrative purposes. The size of the nodes is proportional to their strength, and their color corresponds to module membership. Image courtesy of [19].

coactivation strength was estimated by the Jaccard index, which evaluates the similarities between activation patterns in paired regions over a large number of experiments found in 1,641 t-fcMRI or PET studies, derived from the BrainMap database and conducted between 1985 and 2010 [110]. Four subnetworks, which correspond to four domains, namely, perception, action, emotion, and mixed functions, have been identified by the modularity maximization approach, Figure 1.14. The subnetwork extraction performed on a parallel resting-state functional connectivity network, employing an average group matrix of 27 volunteers, revealed similar subnetworks [19], Figure 1.14. This result confirms that there is strong resemblance, with great spatial overlap, between rest and task subnetworks [111, 112].

Modular brain networks play important roles in conserving wiring costs and creating specialized information and complex dynamics [106]. However, subnetwork extraction remains challenging due to the pronounced noise found in the neuroimaging data. Few methods have exploited intrinsic properties of brain networks other than the modularity.

1.4.3 Multimodal Fusion in Brain Connectivity Analysis

We have introduced the graph theoretical approach to brain connectivity analysis from two perspectives, graph construction and analyzing the constructed graphs. The analysis depends on the accuracy of the graph construction derived from anatomical or functional connectivity estimations. However, the FC estimated from fMRI measurements suffers from a significant number of false positives (paragraph 1.4.2.3.2), and the anatomical connectivity estimated from fiber tract strength, based on dMRI measurements, suffers from a pronounced number of false negative connections (paragraph 1.4.2.3.1). These discrepancies may appear at first sight to be setbacks, but combining both sets of data may in fact help alleviate the problems linked to individual modalities since functional and anatomical connectivity would contain different facets of information.

Human brain studies ranging from healthy brains, neuronal development, and neurological disorders all yield examples of direct correspondence between structural linkages and functional correlations. Such studies indicate that there exists a general structural core of functional networks, sustaining the notion of a global link between anatomical and functional connectivity on a whole-brain scale [105, 113–115]. The underlying mechanism could be that the functional synchronization between spatially distinct brain regions is enabled through neural fiber pathways [116]. Among these studies, statistical analysis demonstrates that the functional connectivity strength is positively correlated with that of anatomical connectivity [114]. However, functional connectivity is also observed between regions where there is little or no anatomical connectivity, which most likely indicates the existence of functional correlations mediated by indirect structural connections, i.e., via a third region [113]. In a sense, the brain structure acts as an anatomical skeleton that condenses the dynamic brain function into a low-dimensional manifold or subspace [117].

The interrelationship between functional and anatomical connectivity revealed by the aforementioned studies indicates the potential benefits of the multimodal fusion of fMRI-dMRI data. Recently, multimodal fusion approaches for *joint* anatomical and functional connectivity estimation and analysis have been explored to better understand the human brain network [85, 118–120]. It has been shown

that more robust parcellation can be produced based on multiple anatomical and functional criteria for the definition of brain regions and their boundaries [54, 121]. Higher inter-subject reproducibility of the connectivity patterns can be achieved by multimodal connectivity inference [85, 119, 120]. Great interest has been seen in multimodal subnetwork extraction, which produces a more stable and biologically meaningful subnetwork structure [6, 85]. In terms of behaviour and population studies, multimodal approaches showed a higher discrimination power between the clinical and control populations than unimodal methods [118]. The details of studies of multimodal neuroimaging data fusion methodologies and their applications in cognitive and clinical neurosciences can be found in review papers [122, 123]. Promising results in these studies confirm the merits of multimodal fusion approaches, which can be applied to both constructing and analyzing the brain network.

1.4.4 Deep Learning in Brain Connectivity Analysis

Deep learning, part of a broader family of machine learning, has recently been proven to be the state-of-the-art approach to enhancing performance in various medical applications [124]. The use of deep learning for neurological applications has only lately become more popular. It started with applying deep learning on standard grid-like brain images for: segmentation [125], tumor or lesion detection [126], registration [127], and classification and prediction of disease stages [128–130]. Most of the work used either unsupervised Deep Auto-Encoder (DAE)/Deep Belief Network (DBN) along with Restricted Boltzmann Machines (RBM) followed by fine tuning in a discriminative setting; or variations of Convolutional Neural Networks (CNN), such as the widely used U-net [131], in a supervised manner. Recent endeavour has been made to discover the brain patterns using graph representation for disease prediction [132–134], which was based on the work by Shuman *et al.* [135] who generalized convolutions to graph structured domains using multiplications in the graph spectral domain. Detailed review of applying deep learning in brain image analysis can be found in [124] and [131].

Very few work has been done in terms of applying deep learning to study the basic components of graph construction and analysis of brain connectome. Re-

searches are still in early stages to discover the nature of brain connectivity based on deep learning. Parcellation has recently been described as a semantic segmentation solved by CNN, but only based on raw sMRI image data [136] or regional cell-body stained images [137]. Deep Auto-Encoder might be a direction to study the brain subnetwork by using the learned weights to reveal the hierarchical relations among regional features [138]. However, the subnetwork results depend on a weight threshold and it is hard to show interpretable subnetworks (relations) [138]. Collectively, the few existing deep learning based brain connectivity studies mostly use only raw brain imaging data, without capturing the topological information from the graph representation, nor utilizing multimodal fusion techniques to combine information from different sources. The common challenge of the aforementioned deep learning work on brain connectivity analysis is the lack of ground truth, which provides a significant amount of manually labeled data for training [124]. Although most of these work used unsupervised feature learning as the first step, the final supervised fine tuning step, which finds the most relevant and essential features for target tasks, may render learned features superficial and misrepresent the complexity of the nature of brain [124] (e.g., in application of Alzheimer's disease prediction, only discriminative patterns related to AD are extracted, which are not necessarily related to healthy brains or other types of neurological disorders). Additional challenges for brain image analysis based on deep learning in general, include:

- (1) The lack of large training data sets: Relatively small amount of available data for brain analysis application (~ 1000 medical images vs. millions for natural images) leads to the related overfitting problems (when learning a large number of parameters in the deep network model).

- (2) Data noise: Neuroimaging data suffer from pronounced noise and image resolution limitation. Recent research showed that current Deep Neural Networks (DNNs) are vulnerable to noise, such as adversarial attacks [139], and low dimension perturbation as small as one pixel attack can greatly affect the performance of the models [140].

- (3) Time consuming acquisition of relevant annotations/labeling [131]. Scarce and expensive medical expertise is needed for high quality annotation of medical imaging data.

(4) Label noise: In computer vision, the noise in the labeling of images is typically relatively low since the labeling is oftentimes intuitive. However, in medical applications, even when data are annotated by domain experts, the consensus is somewhat hard to reach, rendering label noise a significant limiting factor in developing algorithms [131].

(5) Class imbalance: In brain disorder classification and prediction applications, images for the abnormal/disease class are always challenging to acquire. The typical strategy is applying specific data augmentation algorithms to increase the size of the underrepresented class [131].

(6) Deep learning is a black box [124]: It is hard to interpret the features learned and link them to their biological intuition. There has been work trying to understand what intermediate layers of convolutional networks are responding to for natural images [141]. Developing methods to explain the features in brain analysis area is urgent and warrants the further research.

(7) The choice of deep learning architecture/model. It is hard to make decisions when selecting a known architecture (transfer learning), devising a task-specific architecture to train from scratch, or fusing across architectures [131]. Reasons need to be justified to support the decisions.

Due to the aforementioned challenges, it is essential and urgent to learn the basics of the brain network to guide the deep architecture construction. Since a graph is a good representation of the brain, our contributions in brain network analysis based on graph theory can help to guide deep learning based approaches to specific applications, such as disease classification and prediction. We will discuss some promising directions on future deep learning work for the purpose of brain connectivity analysis in the Conclusion (Chapter 6).

1.5 Current Challenges Addressed in This Thesis

The challenges of brain graph construction involve developing methods to define nodes and edges, which represent the adequate information for biological reality [16]. Detecting the boundaries between coherent regions on the basis of any one, or multiple numbers of, the existing criteria remains challenging [117]. On the other hand, the methods for node definition can often be unstable, with small mod-

ifications of the input data leading to large changes in the parcellation results [91]. Devising a reliable method for parcellation is essential. At the same time, one possible cause for this instability might be explained by the lack of explicit noise model [91]. Mainstream connectivity based parcellation depends greatly on the quality of the fMRI or dMRI data, both of which suffer from pronounced noise and image resolution limitation. It is challenging to perform noise reduction and reduce the false positives and negatives in connectivity estimations.

In terms of graph analysis, challenges exist particularly in understanding the segregation and integration of the human brain based on numerous network measures. Important issues currently under development involve improved and domain-appropriate approaches to module/subnetwork detection [106]. Few existing approaches fully utilize informative graphical metric information in subnetwork extractions. At the same time, significant attention should be paid to untangling the contributions of spatial embedding and functional specialization to the definition of network communities [106]. It has been shown that there is no clear one-to-one correspondence between network communities in anatomical and functional connectivity [142]. Hence, it is challenging to apply the promising multimodal fusion techniques without a clear guideline. In line with the complex relationship between different brain connectivities, advanced models and algorithms, e.g., considering higher order relations among network nodes, need to be developed for a joint multimodal analysis of brain networks.

1.6 Problem Statement

Within the framework of studying the human brain connectome as a network based on graph theory, we try to tackle the challenges in the steps of node definition, edge building, network measure estimation, and network analyses (particularly involving subnetwork extraction). As the input of the brain graph, anatomical and functional connectivity analyses provide two views of brain connectomes. The aim of this thesis is to devise novel methods to jointly utilize multimodal information to analyze human brain connectivity. The following Research Questions (RQ) raised concerning the challenges existing in the brain network analysis framework will be addressed in this thesis.

1.6.1 How Can We Improve Parcellation?

- RQ1: Can we devise a more reliable/reproducible method for parcellation?
- RQ2: Can we fuse connectivity information from different imaging modalities to better define the criteria for parcellation?
- RQ3: How to validate improvements on parcellation?

1.6.2 How Can We Achieve Noise Reduction When Constructing Edges?

- RQ4: Can we combat false negatives?
- RQ5: Can we suppress false positives?

1.6.3 How Can We Improve Subnetwork Extraction?

- RQ6: Can we use brain graphical metrics to incorporate more domain-related information?
- RQ7: Can we devise a model which would resemble the brain subnetwork's biological nature?

1.6.4 How Can Multimodal Information Help with Subnetwork Extractions for Brain Connectivity Analysis?

- RQ8: Can we fuse anatomical and functional connectivity to improve subnetwork extraction?
- RQ9: Can we devise a model for brain subnetwork extraction which considers higher order relations among network nodes using multisource information?

1.7 Thesis Contributions

In this thesis, we propose novel approaches, including multimodal fusion techniques, to explore brain connectivity based on the graph theoretical framework.

By constructing the brain graph with a proper definition of nodes and edges, we have achieved improved parcellation and applied noise reduction to perform reliable brain connectivity estimation for the subsequent brain network analysis. We then utilized domain-appropriate graphical metrics to study brain subnetwork extraction. We further propose multimodal fusion approaches towards combining anatomical and functional connectivity for subnetwork extraction, and explore the high order features in brain networks.

1.7.1 Brain Network Node Definition

Our first contribution is towards improving parcellation, i.e., brain network node definition, by utilizing neighbourhood information and multimodal integration techniques.

Specifically, we embedded neighborhood connectivity information into the affinity matrix for the parcellation process to ameliorate the adverse effects of noise, achieving more reproducible parcellation. [P1]

Further, we integrated the connectivity derived from anatomical and functional modalities based on adaptive weighting for improved parcellation. In order to validate these improvements, we designed a number of evaluation metrics including reproducibility, functional homogeneity, leftout data likelihood, and overlaps with cytoarchitectonic areas. [P2-3]

1.7.2 Brain Network Edge Estimation

Our second contribution is to propose noise reduction techniques for brain edge estimation.

We proposed a matrix completion based technique to combat false negatives by recovering missing connections. We validated the effectiveness of this technique using synthetic experiments which simulate the false negatives, and further indirectly verified our technique using an Intelligence Quotient (IQ) prediction application. [P4]

We presented a local thresholding method which can address the regional bias issue when suppressing the false positives in connectivity estimates. We compared this local thresholding method against state-of-the-art thresholding methods

in brain graphs and confirmed the superiority of our approach. [P5]

1.7.3 Graphical Metric Guided Subnetwork Extraction

Our third contribution is to improve brain subnetwork extraction by using multi-pronged graphical metric guided methods. We propose a connection-fingerprint based modularity reinforcement model which reflects the putative modular structure of a brain graph. We show that the subnetworks extracted using our model matched well with well-established brain systems. Compared with methods conducted without using the proposed graphical metric guided strategies, the results of using our approach manifest more biologically meaningful brain systems. [P5]

We also propose a provincial hub guided feedback optimization model, which resembles the brain subnetwork's biological nature for more reproducible subnetwork extraction. The subnetworks derived from this model demonstrates greater overlaps with well-established brain systems as compared to contrasted methods. [P6]

1.7.4 Multimodal Fusion for Subnetwork Extraction

To further improve subnetwork extraction, our fourth contribution is to develop multimodal techniques to integrate information from multiple sources. We first propose a provincial hub guided multimodal random walker based model to fuse anatomical and functional data by propagating the modular structural information across different modalities. [P6]

We next integrate multi-task information into subnetwork extraction based on hypergraph to study the higher order relations among network nodes. [P7] We further propose a co-activated clique based overlapping subnetwork extraction method. [P8]

1.8 Materials and Experimental Setup

1.8.1 Overview of Currently Available Open Access Neuroimaging Datasets

Modern advances in neuroimaging enable systematic exploration of the human connectome. The human connectome research requires massive datasets with minimal variance in experimental procedures acquiring large samples in a reasonable amount of time. Efforts towards building benchmarks with expensive amount of data for datasharing have been seen in the recent years, such as the International Neuroimaging Data-sharing Initiative (INDI) [143], 1000 Functional Connectomes Project [144] and the Human Brain Project [145]. Among those dataset, Human Connectome Project (HCP) is the most applicable dataset for this thesis due to the reasons that we summarize below in Section 1.8.2.

1.8.2 Human Connectome Project

HCP [146] stands out by undertaking a systematic effort to map macroscopic human brain circuits and their relationship to behaviour in a large population of healthy adults using multiple imaging modalities. The HCP consortium provides a dataset acquired from a cohort of 1,200 healthy adults that is made publicly available. The HCP dataset has high spatial and temporal resolution. The typical acquisition resolutions are 2mm isotropic voxel size for dMRI, 4 mm isotropic and repetition time (TR) at two-second for fMRI; however, HCP consortium manages to acquire higher resolution at 1.25 mm for dMRI, 2mm and 0.72 second for fMRI using fast TR sampling. Due to the high quality and extensive scope of the open access HCP data, we have chosen to use HCP for our validation for brain connectivity analysis in this Thesis.

Data from the HCP Quarter three (Q3) were used in this thesis. The Q3 dataset release has data from 80 unrelated healthy subjects. Data from three of the subjects were excluded due to structural abnormalities in the original dataset. We only employed the T1 sMRI, t-fcMRI, rs-fcMRI, and dMRI scans of these 77 subjects (36 males and 41 females, ages ranging from 22 to 35) who have no history of neurological disease. The T1 images have an isotropic voxel size of 0.7 mm. The

rs-fcMRI data comprise two 30 min sessions, acquired at a TR of 0.72 second and an isotropic voxel size of 2 mm. The t-fcMRI data contain seven different tasks (working memory, gambling, motor, language, social cognition, relational processing, and emotional processing), each under 10 minutes at an isotropic voxel size of 2 mm. The dMRI data have an isotropic voxel size of 1.25 mm, three shells ($b = 1000, 2000, 3000s/mm^2$), and 288 gradient directions with six $b = 0$ (B0) images, acquired with right-to-left and left-to-right phase encoding polarities. Further details can be found in [146]. We used the volume parcellation provided in HCP data package [147] to extract grey matter, white matter, and cerebral spinal fluid for each subject. The T1 sMRI volumes were also used for drawing a correspondence between subjects' native space and Montreal Neurological Institute (MNI) space templates. The subject-wise T1 images were registered to MNI space with a FMRIB's Linear Image Registration Tool (FLIRT) 12 degree of freedom (DOF) affine and then a FMRIB's nonlinear Image Registration Tool (FNIRT) nonlinear registration, described in details in [147].

1.8.3 Data Preprocessing

In order to study the brain connectivity, preprocessing of the raw MRI data, and some further estimation need to be performed to obtain the *Adjacency/Connectivity* matrix for both functional and anatomical connectivity estimations.

1.8.3.1 Data Preprocessing for Functional Connectivity

Preprocessing already applied to the HCP fMRI data includes gradient distortion correction, motion correction, spatial normalization to MNI space with nonlinear registration based on a single spline interpolation, and intensity normalization [147]. It was suggested in the HCP preprocessing paper [147] that no slice timing correction needs to be employed, since the fast TR sampling reduces the need for slice timing correction as all slices in each volume are acquired much closer together than in typical fMRI acquisitions ($TR \sim 2.5s$). Additionally, we regressed out motion artifacts, mean white matter and cerebrospinal fluid confounds, and principal components of high variance voxels using compCor [148]. Next, we applied a bandpass filter with cutoff frequencies of 0.01 and 0.1 Hz for rs-fcMRI data.

For t-fMRI data, we performed similar temporal processing, except a high-pass filter at 1/128 Hz was used. The data were further demeaned and normalized by the standard deviation.

Let \mathbf{Z} be a $t \times N$ matrix of preprocessed fMRI time courses, where t is the number of time points and N is the number of voxels of interest. We estimate the FC matrix using Pearson’s correlation: $\mathbf{C} = \mathbf{Z}^T \mathbf{Z} / (t - 1)$.

1.8.3.2 Data Preprocessing for Anatomical Connectivity

Preprocessing already applied to the HCP dMRI data includes B0 intensity normalization, Echo-Planar Imaging (EPI) distortion correction, eddy current correction, gradient nonlinearity correction, and motion artifacts removal [147]. We used the Medical Imaging Interaction Toolkit (MITK) diffusion imaging package [149] for qBall reconstruction and global tractography. Specifically, for qBall reconstruction, we calculated fiber ODF based on CSA [78]. CSA-ODF provides sharp ODFs and exploits multi-shell information, which enables multiple intra-voxel fiber orientations to be more easily resolved. We then performed whole-brain deterministic global tractography on the estimated ODFs using Gibbs tracking [81]. In contrast to the conventional streamline approach, global tractography jointly considers all fiber trajectories in determining the most plausible fiber configuration [81]. This strategy alleviates error propagation along tracts and produces more known tracts missed with conventional approaches.

Fiber count between target brain regions normalized by volume voxels was then computed as an estimate of anatomical connectivity. To account for how there are often too few fibers going through a single voxel, we adopted the strategy in [150] to incorporate information from each voxel’s neighbors, by using a Gaussian kernel at each endpoint and partitioning a tract across spatially proximal brain areas to model endpoint uncertainty. Both ODF estimation and global tractography were done in the subjects’ native space. To draw a correspondence between the native space and the MNI space, we warped the target atlas onto the $b = 0$ volume of each subject before fiber counting.

Chapter 2

Neighbourhood Information Embedding and Multimodal Integration for Improved Parcellation

This chapter focuses on improving parcellation, i.e., brain node definition, which is based on papers [P1-3]. To generate a graph representation of the brain, properly defining the brain nodes is critical. Parcellating the brain through subdivision of the brain into sub-units that are internally homogeneous in certain criteria is a challenging problem. Here, we propose to devise a reliable parcellation method and fuse multimodal information for parcellation. Specifically, we embed neighborhood connectivity information into the affinity matrix to ameliorate the adverse effects of noise [P1]. Meanwhile, we integrate the connectivity derived from anatomical and functional modalities based on adaptive weighting [P2-3].

2.1 Neighborhood Connectivity Informed Parcellation

2.1.1 Related Work

With today’s wide availability of MRI scanners, most strategies are now designed around using only MRI-based features for parcellation, especially connectivity based parcellation [59, 66]. We here focus on unsupervised clustering methods using connectivity data, since there is no ground truth in parcellation applications. Such methods include ICA, Gaussian Mixture Model (GMM) [151], kmeans clustering and its fuzzy alternative [152], hierarchical clustering methods [153], region growing [154], graph-based methods such as Normalized cuts (Ncuts) [155] and community detection method [18], and boundary mapping [156]. The majority of these parcellation approaches rely on calculating an affinity matrix that encodes the connectivity similarity between the image voxels. As such, defining an affinity matrix that accurately represents the underlying structure is central in clustering-based parcellation algorithms.

2.1.2 Neighborhood Information Embedded Affinity Matrix

In order to emphasize on discriminable features in the critical affinity matrix, we have developed a novel approach by embedding neighborhood connectivity information into the affinity matrix. This approach serves the dual purpose of allowing self-adaptive adjustment of voxel affinity values and providing robustness against noise. Our rationale is that, being the sole input to most clustering methods, affinity matrices should encapsulate as much relevant information as possible, rather than just average connectivity similarities. The advantages of our method are two-fold: First, the distinction between voxels at boundaries and those in interior regions, helps preclude pooling information from different parcels, which would result in ambiguous affinity values. Second, such distinction enables emphasizing or de-emphasizing the affinity values based on putative cluster memberships of voxels, ameliorating the adverse effects of noise. The proposed affinity matrix can be used with any parcellation method that takes an affinity matrix as its input.

2.1.2.1 Gaussian Kernel Affinity Matrix

Traditionally, connectivity similarities or distances are mapped using a Gaussian kernel to produce an affinity matrix for clustering. Gaussian kernel function is a typical candidate in cases where the data points live in the Euclidean space while keeping the local information [157]. Thus, we transform the connectivity matrix \mathbf{C} using a Gaussian kernel to better capture connectivity structure as below:

$$\mathbf{A}_{ij} = \exp\left(-\frac{d_{ij}^2}{2\sigma^2}\right), \quad (2.1)$$

where d_{ij} is the distance between voxels i and j , which we set to be $1 - \mathbf{C}_{ij}$, \mathbf{C}_{ij} is the estimate of the connectivity between voxel i and j , and σ is a parameter controlling how rapidly \mathbf{A}_{ij} decays with increasing distance. Usually, σ is set to be the average distance among neighboring voxels over the whole affinity matrix. However, the inherent assumption of fixed density distribution of affinity values rarely holds true in practice. To adaptively tune σ based on the local statistics of voxel neighborhoods, (2.1) has been modified as below in (2.2) as the multiple density kernel, which has been proven to outperform the classical Gaussian kernel [158]:

$$\mathbf{A}_{ij} = \exp\left(-\frac{d_{ij}^2}{\sigma_i\sigma_j}\right), \quad (2.2)$$

where σ_i and σ_j are local scaling parameters. Based on empirical observations, Zelnik *et al.* suggested that setting σ_i to be the distance between voxel i and its 7th nearest neighbor outperforms the fixed density kernel [158].

2.1.2.2 Proposed Affinity Matrix

Although the multiple density kernel has been proven to outperform the classical Gaussian kernel, we argue that the choice of local scaling parameter σ_i is vulnerable to noise. Based on the multiple density affinity matrix, we propose a more robust local density estimation approach and accentuate the difference between intra-cluster and inter-cluster affinity values. Essentially, each voxel in a cluster can be classified as either an Interior Point (IP) or a Boundary Point (BP). An illustration of IP and BP can be found in Figure 2.1. The neighbors of the latter have a heterogeneous cluster membership structure whereas all neighbors of the former

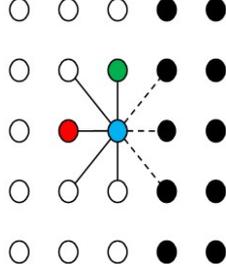


Figure 2.1: Illustration of IP (red dot) and BP (green and blue dots). The neighbors of the latter have a heterogeneous cluster membership structure whereas all neighbors of the former belong to the same cluster.

belong to the same cluster.

Hence, we propose a parameter K indicating voxel's cluster membership by the local neighborhood information defined as below:

$$K_i = L_i - S_i, \quad (2.3)$$

where we denote the distances between voxel i and its neighbors as $\{d_i\}$, the values below the 30th percentile of $\{d_i\}$ as $\{d_i\}_S$, and the values above the 70th percentile of $\{d_i\}$ as $\{d_i\}_L$, L_i and S_i are the median values of $\{d_i\}_L$ and $\{d_i\}_S$. Intuitively, K_i is small when voxel i is an IP, and large when voxel i is a BP, since the spread of $\{d_i\}$ will be larger if voxel i has neighbors from different clusters. With this definition, which indirectly allows us to distinguish between IPs and BPs, we define our affinity matrix as Neighborhood-information-embedded Multiple Density (NMD) kernel:

$$\mathbf{A}_{ij} = \exp\left(-\frac{K_i K_j d_{ij}^2}{\bar{d}_i \bar{d}_j}\right), \quad (2.4)$$

where \bar{d}_i is the average distance between voxel i and its neighbors in the same cluster, which is estimated as the one fourth of its nearest neighbors, replacing the σ_i in (2.2) referring to the distance between voxel i and its 7th nearest neighbor, which is prone to be erroneous in noisy neuroimaging data.

This NMD kernel effectively modifies the \mathbf{A}_{ij} values based on the cluster memberships of voxels i and j . Specifically, it scales \mathbf{A}_{ij} up when i and j are in the

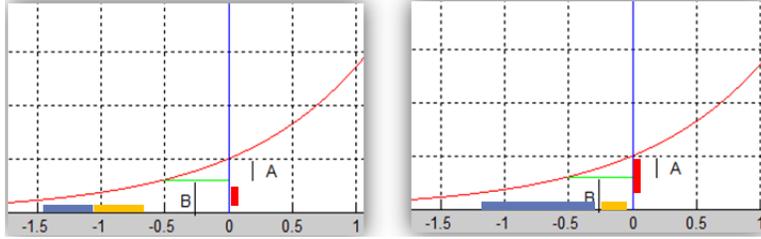


Figure 2.2: Theoretical illustration of NMD affinity matrix. Left is traditional Gaussian kernel, and the right is NMD. The affinity distribution, shown in the red bars, expands to a wider range when NMD is used.

same cluster, and scales it down when they are not. If two neighboring voxels i and j belong to different clusters, they then must both be BPs. In this case, d_{ij} is large and large $K_i K_j$ makes the numerator of the exponential function in (2.4) larger than in (2.2), scaling down \mathbf{A}_{ij} . In noisy cases, an adverse effect of noise is that d_{ij} might decrease. However, K_i and K_j are still large, which suppresses the noise by keeping \mathbf{A}_{ij} relatively small. The opposite holds if i and j belong to the same cluster. In this case, either both i and j are IPs, or one of them is a BP. This will lead to a relatively small $K_i K_j$, scaling up \mathbf{A}_{ij} . In case that noise increases d_{ij} , $K_i K_j$ mitigates the effect of noise by keeping relatively large \mathbf{A}_{ij} . The illustration of how NMD kernel modify the distribution of affinity values for clustering is shown in Figure 2.2.

2.1.3 Experiments and Results

To evaluate our proposed approach, we apply Ncuts to both synthetic and real data from HCP dataset, comparing the self-tuning multiple density [158] in (2.2) with our proposed modified affinity matrix in (2.4). We opt to use Ncuts as the clustering method since it has been shown to outperform GMM. Other clustering algorithms are also applicable, but we have chosen Ncuts for its global optimality guarantees so that our results are not prone to problems, such as instability to initialization in the case of kmeans clustering or obtaining only a local minimum in the case of hierarchical Ward’s clustering. However, it is important to note that our proposed affinity matrix can be used in conjunction with any parcellation method that operates on affinity matrices.

For quantitative validation on synthetic data, we simulated a dataset comprising six horizontally connected clusters on a 3D grid of $5 \times 5 \times 30$ voxels, with each cluster comprising a cubic region of $5 \times 5 \times 5$ voxels. Based on this configuration of voxels, we generated a binary $N \times N$ ground truth affinity matrix, where $N=750$ is the total number of voxels. For each voxel, we simulated time courses by randomly drawing samples from a multivariate normal distribution with zero mean and covariance structure given by the affinity matrix of the data. We added Gaussian noise to the time courses with the SNR set to -10 dB. This noise level was deliberately chosen to be low enough to assess the robustness of our method under severe noise conditions. We repeated this process generating 50 noisy versions of the ground truth.

We then compared the ground truth and the parcellations generated using our approach using Dice Similarity Coefficient (DSC) [159]. DSC is defined as $(2|X \cap Y|)/(|X| + |Y|)$, where X is the index set of voxels in a given parcel from the first parcellation (ground truth here), and Y is the index set of voxels in the matched parcel from the second parcellation (the estimated parcellation here). $X \cap Y$ is the set of voxels commonly assigned to the matched parcel pair. $|\cdot|$ is the cardinality/size of the set. DSC lies between 0 and 1, with 1 indicating the parcel pair comprises exactly the same set of voxels. We perform parcel matching between the two parcellations using Hungarian clustering [160] with DSC between each pair of parcels as the similarity metric. All statistical comparisons are based on the Wilcoxon signed rank test with significance declared at an α of 0.05 with Bonferroni correction.

We show that higher clustering accuracy quantified by the average DSC can be attained with our modification of the multiple density affinity matrix, with statistically significant improvement from 0.90 ± 0.038 to 0.94 ± 0.040 .

On the real data, cortical parcellation maps are challenging to validate due to the lack of ground truth. However, assuming that there truly is a functional parcellation, it should presumably remain stable for each subject and consistent across subjects [161]. Thus, we based our quantitative validation on intra-subject test-retest reliability (parcellations generated from two time courses of fMRI data of the same subject), and inter-subject test-retest reliability (parcellations from two different subjects) of the resulting brain parcels, which we measured using DSC

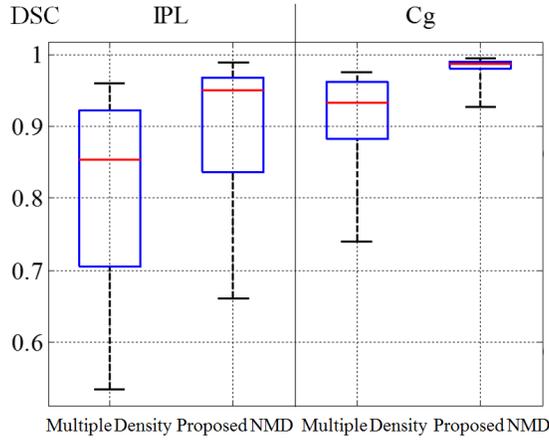


Figure 2.3: Inter-subject test-retest reliability quantified using DSC in IPL and Cg for the HCP data. As reflected in the box plots, our method consistently outperforms the state-of-art method. The blue rectangle spans the first quartile to third quartile. The red line indicates the median and the black whiskers indicate the minimum and maximum.

after relabelling parcels using the Hungarian algorithm [162] to match labels. The Inferior Parietal Lobule (IPL) and Cingulate Cortex (Cg) have been used for validation, since they possess great functional and anatomical heterogeneity. Here, we only validated on fMRI data and set negative connectivity to zero before generating the affinity matrix, since there is a lack of anatomical evidence for negative correlations unlike positive correlations, mentioned in Section 1.8. We show that our method exhibits statistically significantly higher intra-subject test-retest reliability from 0.90 ± 0.04 to 0.95 ± 0.03 for IPL and from 0.93 ± 0.02 to 0.98 ± 0.01 for Cg. Our approach also achieved statistically significant improved inter-subject test-retest reliability with corresponding boxplot shown in Figure 2.3.

Further, we qualitatively evaluate our method by visually comparing the group parcellation map with those estimated from individual subjects. Our rationale here is that the group map is expected to be more reliable since it is generated by pooling data across multiple subjects, effectively increasing the SNR. It can thus be used as pseudo ground truth to compare individual parcellation maps against. For brevity, we only present subjects showing the highest and lowest agreement with

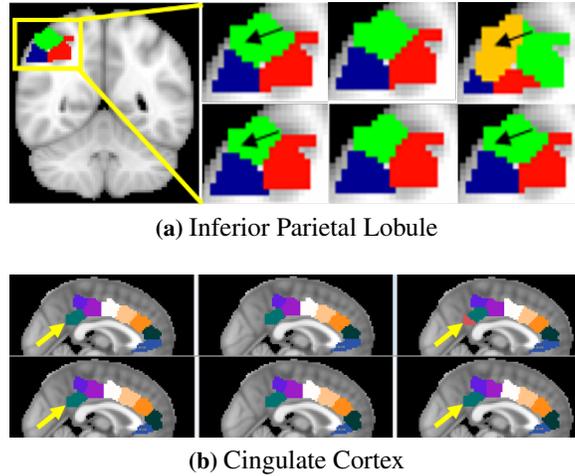


Figure 2.4: Qualitative parcellation results for IPL and Cg using the HCP data. In each area - Top: Using the multiple density in (2.2); Bottom: Using our proposed affinity matrix NMD in (2.4). The 1st column shows the group maps while the 2nd and 3rd columns are the subject-specific maps with the highest and lowest DSC with group maps, respectively. The significant differences between the subject-specific map having the lowest DSC (right) and the group map (left) are highlighted with arrows.

the group map as measured by DSC in Figure 2.4. The group maps obtained using (2.2) and proposed affinity matrices using (2.4) did not exhibit major differences, implying that using multiple density [158] for parcellation suffices when there is enough data. However, the subject-specific map having the lowest DSC (top right) with the group map shows major differences compared to the group map (top left) when (2.2) is used. Specifically, both the number of parcels in the slice shown and the parcel boundaries are significantly different. In contrast, using our proposed affinity matrix results in much more consistent results (bottom left versus bottom right).

2.1.4 Discussion

We proposed a novel affinity matrix estimation for brain parcellation based on multiple density kernel distribution. Our approach can be used in conjunction with any parcellation method that takes an affinity matrix as its input. On synthetic data,

we demonstrated that our proposed affinity matrix (NMD) in equation (2.4) can better represent the data structure by capturing neighborhood connectivity leading to more accurate results compared to the multiple density affinity matrix in equation (2.2). On real data from HCP subjects, we demonstrated the superiority of our method in terms of better intra-subject test-retest reliability and higher inter-subject test-retest reliability. Qualitatively, we showed that subject-specific parcellation maps better resemble the group maps when parcellating using our affinity matrix. Our next step focuses on the extension to whole-brain parcellation and devising a multimodal integration approach.

2.2 Multimodal Connectivity Fusion for Parcellation

2.2.1 Related Work

The majority of recent studies focuses on using fMRI data for parcellation [154, 155, 163, 164]. A general limitation of these unimodal approaches is that they do not capture other brain attributes, such as fiber pathways (serving as the physical substrate for functional interactions), which provide additional indication of how the brain is organized. Incorporating this information would presumably provide more reliable parcellation than using fMRI data alone. Towards this end, consensus maps have been generated from overlaps among probabilistic parcellation maps which were derived from different modalities [165]. To the best of our knowledge, we are the first to propose integration of multimodal information for brain parcellation in 2015 [2], and we recently extended our strategy in [3]. Cortical surface parcellation approaches based on multimodal information have recently been proposed [54, 121]. We note that our multimodal volumetric parcellations could serve as excellent complements to these two recent multimodal cortical surface parcellation [54, 121], especially considering the importance of subcortical regions for clinical investigations, such as Parkinson’s disease studies.

2.2.2 Proposed Multimodal Integrated Parcellation Framework

Most existing parcellation methods use structural information (e.g., gyri and sulci), FC, or AC alone to group voxels into parcels. Each of these modalities has its own

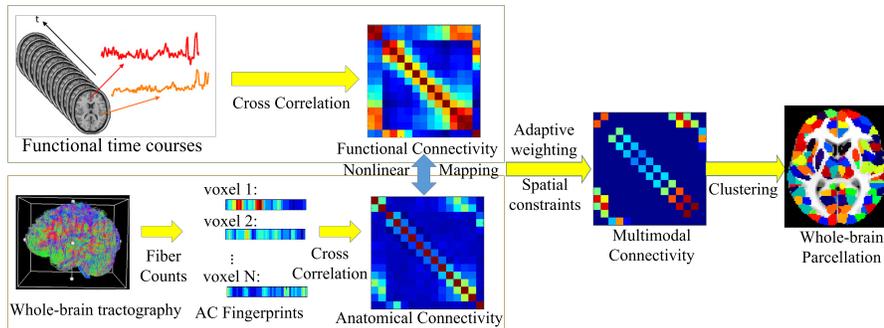


Figure 2.5: Flowchart of proposed approach for multimodal brain parcellation. AC and FC estimates are generated for each voxel to derive voxel-by-voxel connectivity matrices followed by a distribution normalization. The resulting matrices are integrated to generate multimodal connectivity estimates. Ncuts is applied on the integrated similarity matrices, to generate whole-brain parcellation.

inherent limitations. For instance, structural attributes might not relate to brain function, FC estimates are prone to false positives, and AC estimates are prone to false negatives. We thus examine whether combining information across these brain attributes could alleviate the limitations of each modality. To prevent bias towards any particular task, we use Resting State Functional Connectivity based on MRI data, which captures intrinsic functional connectivity. At the same time, we use dMRI data to estimate AC. In order to incorporate structural information on gyri and sulci, we present a region level extension of our approach.

We propose an approach for integrating multimodal information, as summarized in the following and the corresponding workflow is shown in Figure 2.5.

- Generate AC and FC estimates for each voxel to derive voxel-by-voxel connectivity matrices.
- Map AC and FC estimates to a common value range using a distribution normalization function.
- Integrate these matrices based on adaptive weighting, where the weights are derived from the reliability of the AC and FC estimates.

- Apply standard clustering algorithms, such as Ncuts, on the integrated similarity matrices, to parcellate the brain.

Our multimodal approach can be used for both subject and group level parcellation. For subject parcellation, we derive the integrated similarity matrix (derived from the connectivity matrix) from each subject’s AC and FC estimates. For group level parcellation, we average the AC and FC estimates across subjects prior to AC-FC integration.

2.2.3 Anatomical and Functional Connectivity Estimation

Our AC estimation is based on a fingerprint concept [65], that the voxels having similar fiber connection fingerprints should have higher connectivity strength. This type of connectivity estimation can incorporate both node-to-node relationship and the fingerprint pattern reflecting the relationship between one node and the remaining regions.

Given the estimated fiber tracts (derived by following the steps in Section 1.8), we first define an AC fingerprint for each voxel as the number of tracts connecting that voxel to a set of target regions. Target regions can be taken as every other voxels or region of interest from existing atlases. We use the 112 regions in the well-established HO atlas, because it consists of the highest number of subjects with both manual and automatic labelling technique compared to other commonly used anatomical atlases [166]. AC is then estimated as the cross-correlation between the AC fingerprints of each voxel pair, the resulting matrix is denoted as C^A . Using the HO atlas as target regions has multiple advantages. First, this strategy reduces the number of unknowns to be estimated, i.e., the number of elements in each AC fingerprint is reduced from $>100,000$ to 112. Second, using ROIs instead of voxels provides more tolerance for uncertainty in fiber endpoint location [13]. Moreover, this strategy lowers computational cost.

We note that we spatially normalize the HO atlas to the subject’s native space to avoid distorting the dMRI volumes, which impacts the qBall estimation. To constrain the parcels to be spatially-contiguous, we keep only AC estimates of the 26 neighbors for each voxel, which include voxels of one-ring neighbors in Cartesian coordinates [164, 167]. The spatial contiguity within estimated parcels ensures

the representation of anatomically homogeneous regions, and hence preserve the interpretability of the connectivity results [168]. Additionally, the spatial contiguity distinguish network nodes from large-scale networks of nodes, i.e., subnetwork extraction [168].

On the other hand, the FC estimates, denoted as \mathbf{C}^F , are derived by following the steps in Section 1.8. Let \mathbf{Z} be a $t \times N$ matrix of preprocessed rs-fcMRI time courses, where t is the number of time points and N is the number of voxels of interest. We estimate the FC matrix using Pearson’s correlation: $\mathbf{C}^F = \mathbf{Z}^T \mathbf{Z} / (t - 1)$. Also, only \mathbf{C}^F of the 26 neighbors of each voxel are retained to enforce spatial continuity.

For our estimation of both the AC and FC, we set negative values in the both connectivity matrices to zero due to the difficulty in interpreting negative connectivity (there is a lack of anatomical evidence for negative correlations unlike positive correlations) [93].

2.2.4 Distribution Normalization

It is important to note that naive combination of AC and FC values may not be suitable, since their estimates have different distributions and there is a lack of straightforward correspondence between these two estimates, Figure 2.6. Take value 0.5 for example, which is in the overlapping area in Figure 2.6, it refers to high correlation in the FC distribution, but low correlation in the AC distribution, leading to a misleading combined correlation value. This difference is important to take into account to eliminate the risk of overweighting one modality over the other [2].

To ensure no possible overweighting in generating combined connectivity estimate occurs, we apply a nonlinear mapping based on distribution normalization to gain a better correspondence between the two modalities. We use histogram matching to match the distribution of AC values to that of FC [2].

Let $a_1 \leq a_2, \dots, \leq a_M$ be the raw AC values, and $f_1 \leq f_2, \dots, \leq f_M$ be the FC values, where M is the number of voxel pairs, we compute the histograms of AC and FC values, as well as their cumulative distribution functions $F_a(x) = P(X \leq x), x \in [a_1, a_M]$ and $F_f(y) = P(Y \leq y), y \in [f_1, f_M]$. $P(\cdot)$ is the probability that the

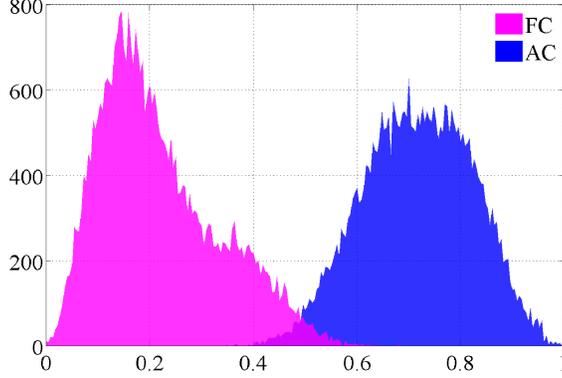


Figure 2.6: Histograms of FC and AC derived from the group data of 77 HCP subjects. The difference between the two distributions could potentially cause bias if the FC and AC are naively fused.

random variable X takes on a value less than or equal to x . Next, we replace each AC value a_i with the FC value f_j , which satisfies $F_a(a_i) = F_f(f_j)$ as:

$$\phi(a | F_a(a) = F_f(f)) = f. \quad (2.5)$$

This distribution normalization procedure enables unbiased integration of AC and FC, where each mapped AC value is used in the multimodal connectivity estimation in the following section.

2.2.5 Multimodal Connectivity Estimation using Adaptive Weighting

To overcome the inherent limitations associated with each of the two modalities, we propose combining connectivity estimated from two modalities using an adaptive weighting scheme based on voxel-wise reliability:

$$\mathbf{C}_{ij}^M = (1 - \mathbf{w}_{ij})\mathbf{C}_{ij}^F + \mathbf{w}_{ij}\phi(\mathbf{C}_{ij}^A), \quad (2.6)$$

where \mathbf{C}_{ij}^M is the Multimodal Connectivity (MC) estimates between voxels i and j , \mathbf{C}_{ij}^A and \mathbf{C}_{ij}^F are AC and FC estimates, and \mathbf{w}_{ij} is the voxel-wise adaptive weight.

The adaptive weighting which enables us to integrate multimodal connectivity estimates, affectively weighs down the contribution of FC in voxels where fMRI

observations are deemed to be unreliable and AC estimates are reliable. We define the integration weights based on a connection reliability ρ as:

$$\mathbf{w}_{ij} = (1 - \rho_{ij}^F) \rho_{ij}^A, \quad (2.7)$$

where we construct a measure of connection reliability ρ based on the local voxel reliability γ of the FC and AC estimates defined as follows. For FC, we split the time course of each voxel i into two halves and compute the Pearson’s correlation between its local fingerprints (i.e., similarity between a voxel and its 26 neighbors) derived from these two halves, which we take as the FC voxel reliability γ [2]. For AC, we extract two subsets of dMRI data by downsampling the single qBall acquisitions by minimum distance pair matching (see **Reproducibility** in Section 2.2.7). We then generate local fingerprints for the two subsets and use their Pearson’s correlation as the AC voxel reliability. Note that we use the information from the immediate 26 one-ring neighbors of a given voxel, as it has been shown that larger neighborhoods might make the analysis more prone to acquisition and registration artifacts [114]. The reliability of an edge connection is implicitly bounded by one of the two end voxels having relatively lower reliability. As such, we define ρ_{ij} as the minimum reliability of edge end voxels i and j as:

$$\rho_{ij} = \min(\gamma_i, \gamma_j) \quad (2.8)$$

where γ_i is the reliability of voxel i , with γ_i derived from either FC or AC.

2.2.6 Clustering and Region Level Extension

We apply Ncuts on \mathbf{A} to generate brain parcellation. We opt to use Ncuts for its global optimality guarantees to avoid instability to initialization and local minimum as mentioned in Section 2.1.3. With regard to the number of parcels, there are indeed data-driven strategies to select the number of parcels, e.g. based on predictability on leftout data [68] and stability over data splits [69]. However, these criteria tend to return different numbers of parcels depending on the number of samples (time samples/subjects) available. Also, the number of parcels largely depends on the application and ease of result interpretation. Hence, instead of finding

an “optimal” number, we adopt a multi-scale perspective by examining a range of values 256, 512, 1024. 256 parcels is examined to compare with existing anatomical atlases [166], which have between 100 to 200 regions. 512 parcels is selected based on prior studies on architectonic [70] and functional [71] information that suggest the human brain comprises ~ 500 regions. 1024 parcels is chosen since finely grained brain atlas e.g. with ~ 1000 parcels has also been used in well-established studies [114].

We further present a simple region level extension of our parcellation approach that substantially reduces computational cost and simplifies the problem of parcel correspondence between subjects, while having a side-advantage of incorporating structural information on gyri and sulci. Specifically, we decompose the whole-brain parcellation problem into multiple ROI parcellation subproblems, by separately parcellating each ROI of the HO atlas (which was derived based on sMRI information). Exactly the same procedure used in whole-brain parcellation can be applied at region level. The number of parcels within each ROI is set based on the proportion of voxels in the brain covered by the given ROI.

2.2.7 Design of Evaluation Metrics

2.2.7.1 Reproducibility

Based on the assumption that a parcellation exists, the parcellation should presumably remain stable for each subject and consistent across subjects [161], we evaluate the parcellations on three types of reproducibility, namely intra-subject test-retest reliability, inter-subject test-retest reliability, and subject-group consistency.

To measure intra-subject test-retest reliability, we first divide the rs-fcMRI data and dMRI data into two splits and generate two subject-specific parcellations from each data split. For functional data, we use two rs-fcMRI scans accessible from HCP dataset to assess reproducibility. In terms of diffusion data, we propose a method for splitting a dMRI dataset into two subsets to assess reproducibility. Given a dMRI dataset from a uniformly distributed qBall sampling scheme, we first find all pairs of closest orientations based on minimum distance. We then use

local greedy search to assign each element i of an orientation pair to the subset with elements farthest away from element i . Iterating this procedure generates two exclusive sets of approximately uniformly distributed orientations. We then compare the two parcellations using DSC, as defined in Section 2.1.3. The DSC of all matched parcel pairs are then averaged and taken as the intra-subject test-retest reliability.

For inter-subject test-retest reliability, exactly the same procedures are performed except we compare group parcellations generated from random half splits of the subjects. 10 random splits are performed.

Moreover, we assess the subject-group consistency with the same procedures by comparing each subject level parcellation and the group level parcellation (generated from all subjects' data).

Furthermore, we assess the stability of the group parcellations with respect to the number of subjects, N_s , with N_s set between 5 and 75 at 5 subjects increment. The DSC between the group parcellations built with all subjects vs. N_s subjects is used to estimate stability.

2.2.7.2 Functional Homogeneity

To assess functional homogeneity, we employ two metrics, namely connectivity homogeneity [163] and temporal homogeneity [155]. The first metric evaluates how much variance in the parcel FC patterns is captured by the largest principal component (FC pc ratio). If all voxels within a parcel have similar FC patterns, then the largest principal component would capture a large proportion of the variance. The other metric (mean distance) evaluates the homogeneity of temporal patterns based on the Euclidean distance between the fMRI time courses of each voxel within a parcel and the mean time courses across all voxels within that parcel. If all voxels have similar temporal patterns, then their Euclidean distance from the mean should be small. For both metrics, the averages over parcels are reported.

2.2.7.3 Leftout Data Likelihood

Accuracy of connectome estimation heavily depends on brain parcellation. Thus, another way to evaluate the parcellations is to assess the generalizability of con-

nectivity estimates derived from one scan to an unseen scan. For this, we employ the leftout log data likelihood [169]. Given a parcellation generated from the first scan (using both rs-fcMRI and dMRI data for multimodal approach), we generate parcel time courses for each subject by averaging the voxel time courses within each parcel. Assuming a multivariate Gaussian distribution, we then estimate the inverse covariance, \mathbf{K} , from the first scan and the sample covariance, \mathbf{S} , from the second scan. The leftout log data likelihood is given by:

$$\log\left(\frac{1}{\sqrt[n]{2\pi\det(\mathbf{K})}} \exp\left(-\frac{\text{tr}(\mathbf{SK})}{2}\right)\right), \quad (2.9)$$

where $\det(\cdot)$ is the determinant, $\text{tr}(\cdot)$ is the trace, and n is the number of samples used to compute \mathbf{S} . The role of the first and second scans were then switched, with the average of the two likelihood estimates reported.

2.2.7.4 Overlaps with Cytoarchitecture

We examine the overlaps between the boundaries of our parcels and the areal borders of the Juelich probabilistic cytoarchitectonic mapping thresholded at 0.25 [170]. Parcels with 70% of its volume lying within a given cytoarchitectonic area are considered as part of that area. To avoid false boundary alignments arising from the structural constraints of the HO atlas, we assess overlaps between the cytoarchitectonic areas and the multimodal group parcellation generated without using the region level extension.

2.2.7.5 Subnetwork Extraction

To attach a functional meaning to the parcels, we group the parcels into subnetworks and correspond the subnetworks to established brain systems [18]. Specifically, we first compute the Pearson’s correlation matrix between the rs-fcMRI parcel time courses and apply a threshold of 0.2, which helps reduce the influence from non-significant connections [171]. Ncuts is then applied to group the parcels. To set the number of subnetworks, we search from 5 to 45 at a step size of 1, and choose the value that maximizes modularity. The optimal number of subnetworks found with this scheme is 9. We also examine setting the number of subnetworks

to 25 following previous studies [18]. We further use the same procedures of inter-subject reproducibility calculation described in Section **Reproducibility** to assess consistency of subnetworks generated from two exclusive datasets (two subsets of Q3 from HCP). We apply Hungarian clustering to match the two sets of subnetworks with DSC between node labels subnetworks used as the similarity metric.

2.2.8 Experiments and Results

We evaluated the proposed multimodal parcellation approach based on reproducibility, functional homogeneity, leftout data likelihood, and cytoarchitectonic overlap. We further grouped the parcels into subnetworks and associated them to well established brain systems [18]. Multimodal parcellations at various resolutions derived using the whole-brain and region level schemes as well as parcels' subnetwork labels are provided in the supplemental materials and (will be made) available online. Note that we empirically set w_{ij} to 0 when ρ_{ij} derived from AC dropped below 0.5 to only incorporate reliable AC. For comparisons against our multimodal approach (MC), we examined unimodal parcellations based on FC, AC, and physical distance (DC), random parcellation (RM) with roughly equal-sized parcels [53], as well as other existing atlases based on functional information [155, 163, 164] and anatomical information [57, 172] (HO and Automated Anatomical Labeling (AAL) atlases). Unless stated otherwise, all results are based on parcellations generated using the region level extension for both unimodal and multimodal strategies with the number of parcels set to 512. The parcellation derived based on our multimodal approach at the 512 scale at the group level has been shown in Figure 2.7 a. All statistical comparisons are based on the Wilcoxon signed rank test with significance declared at an α of 0.05 with Bonferroni correction.

2.2.8.1 Reproducibility

To assess intra-subject test-retest reliability, we first generated two brain parcellations for each subject from the two rs-fcMRI scans and two subsets of the dMRI volumes using our multimodal approach. The average DSC across parcels between the two parcellations was used to estimate test-retest reliability. We compared multimodal parcellations against those generated solely based on fMRI data.

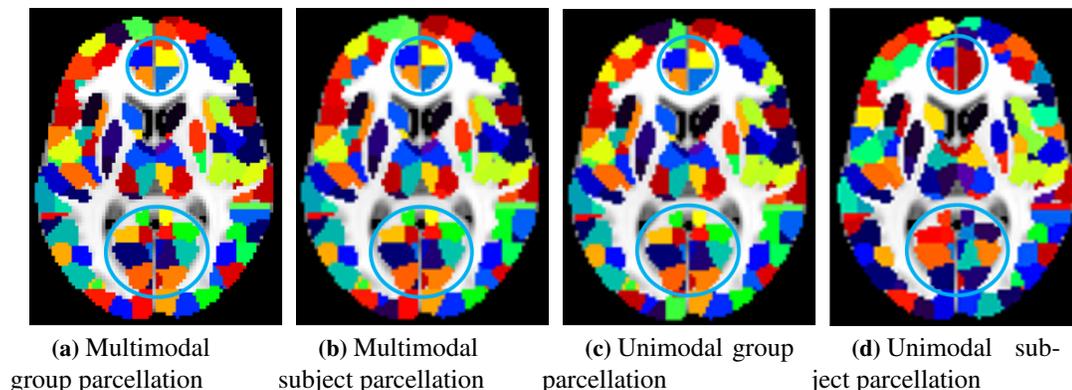


Figure 2.7: Group-wise parcellation and Subject-group consistency. Both the parcellations based on the proposed multimodal approach (left two) and unimodal FC based parcellations (right two) are displayed at 512-parcel scale in 2D space. Second and fourth column: Subject level parcellation having the lowest DSC with the group parcellation. The proposed multimodal parcellations attain higher subject-group consistency (a & b), compared to unimodal FC based parcellations (c & d), which we highlight certain areas using blue circles.

Multimodal parcellations achieved a DSC of 0.949 ± 0.009 , which is significantly higher ($p = 0.00475$) than that generated with the unimodal FC approach (0.939 ± 0.011). The same trend was observed with the number of parcels set to 256 (MC: 0.980 ± 0.0078 against FC: 0.978 ± 0.0078 at $p = 0.0017$), and 1024 (MC: 0.9091 ± 0.01 against FC: 0.9040 ± 0.01 at $p = 1.46e-7$), Figure 2.8.

We next estimated inter-subject test-retest reliability of the group parcellations by first randomly splitting the 77 subjects into 2 halves in 10 different ways. For each random subject split, we computed average DSC across parcels between the pair of group parcellations. Multimodal group parcellations achieved a significantly higher DSC of 0.9507 ± 0.0049 ($p = 0.0098$) compared to that of FC based parcellations (0.9419 ± 0.0053). The trend persisted with the number of parcels set to 256 (MC: 0.9906 ± 0.0022 against FC: 0.9799 ± 0.0030 at $p = 0.0020$), and 1024 (MC: 0.9336 ± 0.0026 against FC: 0.9178 ± 0.0038 at $p = 0.0019$). We observed the difference of inter-subject test-retest reliability in different brain regions, Figure 2.9.

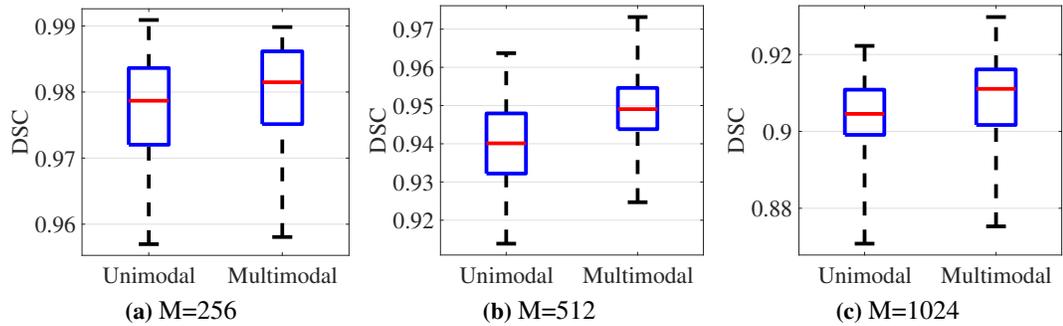


Figure 2.8: Intra-subject test-retest reliability using the DSC as the evaluation criterion. The number of parcels M was set to 256, 512 and 1024, respectively. The blue rectangle spans the first quartile to third quartile. The red line indicates the median and the black whiskers indicate the minimum and maximum. Multimodal parcellations achieved significantly higher test-retest reliability than those generated based on fMRI data alone.

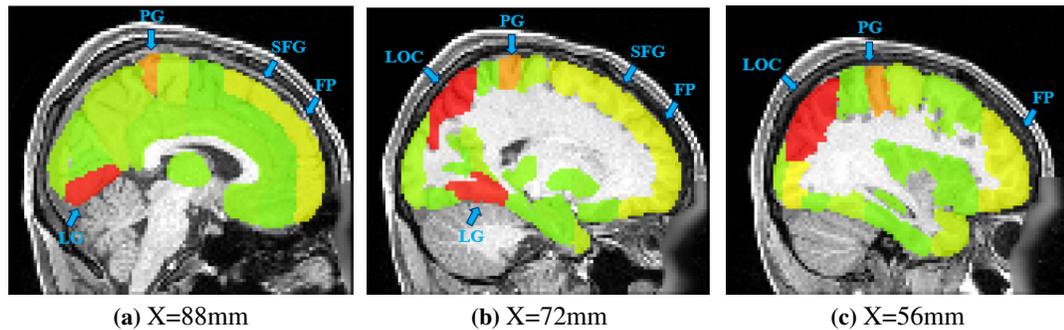
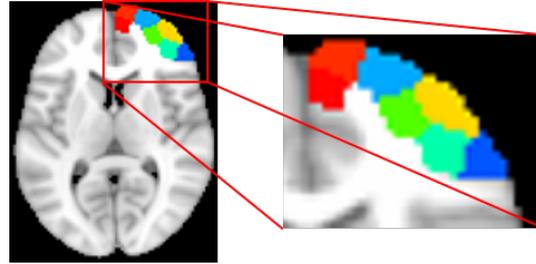
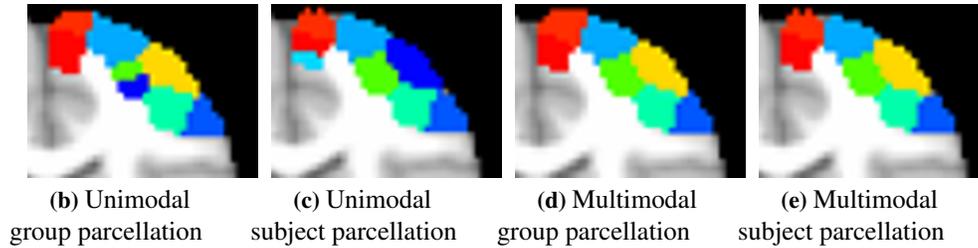


Figure 2.9: The inter-subject test-retest reliability (range from low red to high green) of each brain region of the Harvard-Oxford atlas (only shown on sagittal slices $X=88\text{mm}$, 72mm and 56mm as exemplars). Regions, such as the Lingual Gyrus (LG), Lateral Occipital Cortex (LOC), Postcentral Gyrus (PG), Frontal Pole (FP), and Superior Frontal Gyrus (SFG) show relatively lower reproducibility.



(a) Frontal pole (right) illustrated as an inset within the whole brain (left)



(b) Unimodal group parcellation (c) Unimodal subject parcellation (d) Multimodal group parcellation (e) Multimodal subject parcellation

Figure 2.10: Subject-group consistency. Comparison between the group parcellation and the subject parcellation of frontal pole having the lowest DSC with the group parcellations.

We further evaluated the consistency between subject level and group level parcellations, described in **Reproducibility** in Section 2.2.7. Quantitatively, multimodal approach significantly improved the subject-group consistency from 0.872 ± 0.0099 for the unimodal approach to 0.883 ± 0.0105 ($p = 0.00138$) at the scale of 512 parcels. We observed similar trend with the number of parcels set to 256 (MC: 0.8876 ± 0.0127 against FC: 0.8826 ± 0.0129 at $p = 2.46 \text{ e-}14$), and 1024 (MC: 0.5098 ± 0.011 against FC: 0.5069 ± 0.011 at $p = 1.523 \text{ e-}13$).

Qualitatively, we compared the subject parcellation having the lowest similarity with the group parcellation based on Dice coefficient (the worst case). We observed that multimodal parcellations have higher subject-group consistency than its unimodal counterpart (Figure 2.7). To more clearly illustrate the differences, we plotted the parcellations of an exemplar ROI, namely the frontal pole of the right hemisphere in Figure 2.10. Major differences between the subject (Figure 2.10 c) and the group parcellations (Figure 2.10 b) were observed for the unimodal ap-

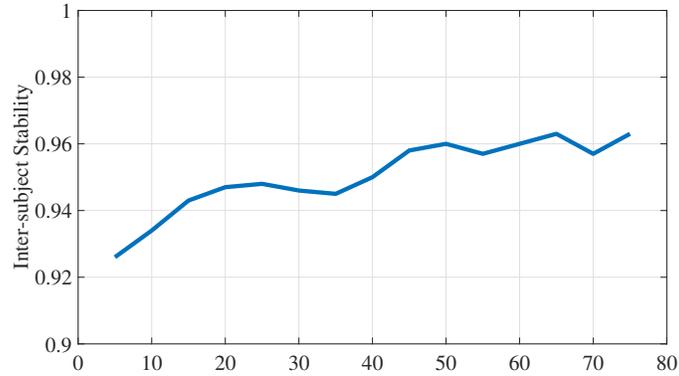


Figure 2.11: Stability with respect to the number of subjects in generating group parcellations. We compared the multimodal group parcellations generated from all 77 subjects against N_S subjects with N_S set between 5 and 75 at interval of 5 subjects. Stability was estimated using the Dice coefficient. The number of parcels was set to 512. The stabilities of the group parcellations as measured using the Dice coefficient plateaued at 0.96 after $N_S = 50$.

proach in both the number of parcels and the parcel boundaries, whereas substantially more consistent parcels were found with multimodal approach (Figure 2.10 d & e).

Moreover, we examined the stability of the group parcellations with respect to the numbers of subjects. Specifically, we compared the multimodal group parcellations generated from all 77 subjects against using only N_S subjects with N_S set between 5 and 75 at interval of 5 subjects. The stabilities of the group parcellations as measured using the DSC plateaued at 0.96 after $N_S = 50$, Figure 2.11.

2.2.8.2 Homogeneity

We assessed the functional homogeneity using the FC pc ratio and the mean distance, defined in Section 2.2.7 **Functional Homogeneity**. We compared our multimodal parcellations against unimodal parcellations derived from AC, FC, DC, and RM. Multimodal parcels achieved significantly higher functional homogeneity in both FC pc ratio and mean distance (0.3751 and 67.2022) than those based on AC (0.3681 and 67.4321), DC (0.3695 and 67.6756), and RM (0.3461 and 68.1732),

and comparable to parcels based on FC (0.3750 and 67.2028), Figure 2.12. Similar trend was observed with the number of parcels set to 256 and 1024. When the number of parcels was set to 256, multimodal parcels achieved significantly higher functional homogeneity in both FC pc ratio and mean distance (0.3493 and 67.0495), than AC (0.3482 and 67.1005), DC (0.3393 and 67.4590), and RM (0.3335 and 67.4682) at $p < 0.05$ based on the Wilcoxon signed rank test, and comparable to parcels based on FC (0.3493 and 67.0497). When the number of parcels was set to 1024, multimodal parcels achieved significantly higher functional homogeneity in terms of both FC pc ratio and mean distance (0.3942 and 66.7457) than FC (0.3938 and 66.7591), AC (0.3936 and 66.7536), DC (0.3937 and 66.7883), and RM (0.3819 and 67.0490) at $p < 0.05$. To compare against existing atlases [57, 155, 163, 164, 172], we also generated parcellations with comparable number of parcels. Details on existing anatomical and functional atlas can be found in Table 2.1 and Figure 2.13. We note that to compare against the AAL and HO atlases, which has about a hundred parcels, we could not use the region level strategy, i.e., dividing regions of the these two atlases will obviously result in more regions than the original one. We thus instead used the whole-brain parcellation approach without the region level strategy for comparison against AAL and HO atlases. Multimodal parcellations achieved significantly higher functional homogeneity compared to existing atlases except for Gordon’s atlas [163] with $p = 0.2372$ (Figure 2.14), whose higher functional homogeneity is likely due to the bias arising from parcellating only a thin shell of the cortex (see Figure 2.13). Specifically, grouping fewer voxels into the same number of parcels would result in smaller parcels, which inherently would have higher functional homogeneity than larger parcels (i.e., more voxels would introduce more variability).

2.2.8.3 Leftout Data Likelihood

We further evaluated the multimodal parcellations with leftout log-data likelihood. The mean leftout log data likelihood of multimodal parcellations was higher than that of FC and AC though not statistically significant ($p = 0.53$ and $p = 0.37$). RM and DC attained higher leftout log-data likelihood than AC, FC and MC. The trend moderately changed with the number of parcels set to 256 and 1024. Specifically,

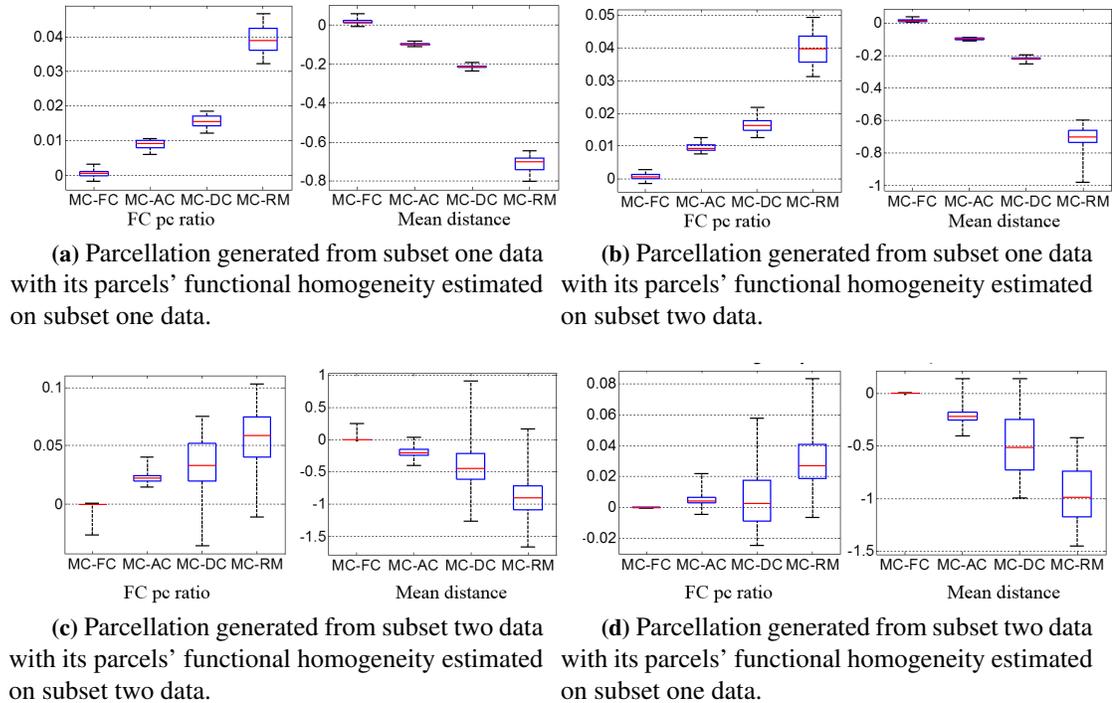


Figure 2.12: Functional homogeneity comparisons between multimodal and unimodal parcellations. In order to show consistent gain in each case of comparisons, plots of MC-FC, MC-AC, MC-DC and MC-RM are provided. Positive FC pc ratio and negative mean distance indicate higher functional homogeneity gained from multimodal parcellations. Number of parcels were set to 512. Multimodal parcels achieved significantly higher functional homogeneity in both FC pc ratio and mean distance, see text for details.

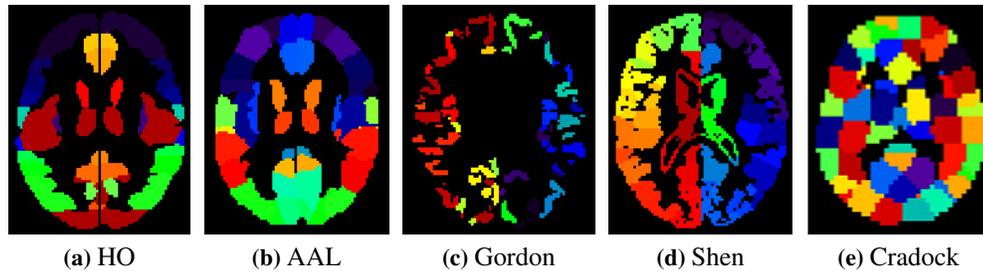
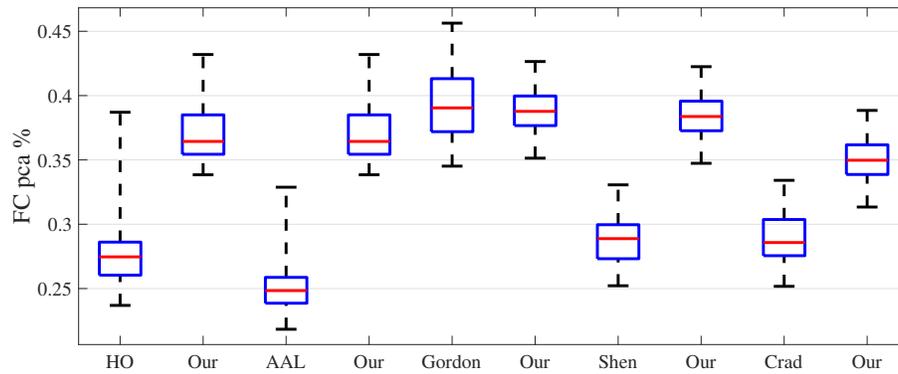
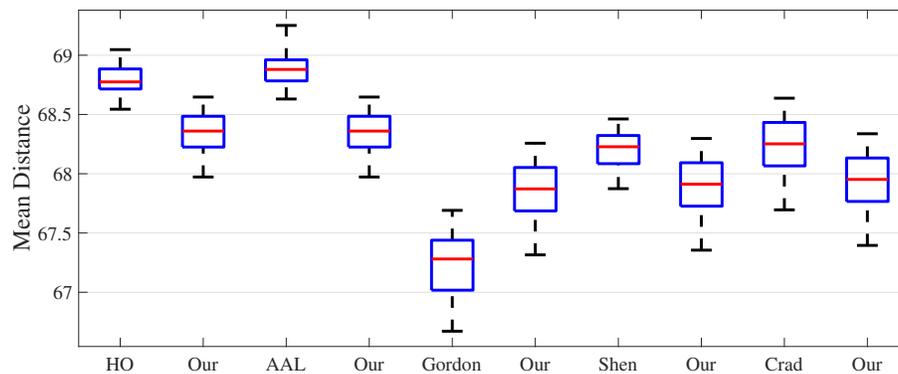


Figure 2.13: Existing anatomical and functional parcellations



(a) FC pc ratio



(b) Mean distance

Figure 2.14: Functional homogeneity comparison between multimodal parcellations and existing atlases. Number of parcels for multimodal parcellation was set to the exact number of parcels in the contrasted atlas. Contrasted atlases included HO (112 parcels), AAL (116 parcels), Gordon's (333 parcels), Shen's (268 parcels), and Craddock (190 parcels). Multimodal approach achieved significantly higher homogeneity against the contrasted atlases except for Gordon's.

Table 2.1: Details on existing anatomical and functional parcellations.

Name	Reference	# of parcels	Notes
HO	[57]	112	Volumetric format
AAL	[172]	116	Volumetric format
Gordon	[163]	333	Surface to volume mapping
Shen	[155]	268	Volumetric format
Craddock	[164]	190	Volumetric format

at the scale of 256 parcels, the likelihood of MC was higher than FC and DC, and significantly higher than AC. At the scale of 1024 parcels, the likelihood of MC was significantly higher than FC, AC, and DC (Figure 2.15). Here, RM and DC at some scales perform better due to the regular sampling of volume of interest from parcels with roughly equal size [169].

2.2.8.4 Cytoarchitecture

We examined whether multimodal parcellations (without applying the region level extension) match cytoarchitectures by comparing them to the areal borders of the Juelich probabilistic cytoarchitectonic atlas (thresholded at 0.25) [170], Figure 2.16. We opted to use parcellations generated without applying the region level strategy for this analysis since regional boundaries in the HO atlas might be inherently aligned to some of the cytoarchitectonic areas, hence positively biased. For clearer visualization, we mapped the multimodal parcellations from MNI space to the PALS midthickness surface using the Caret software [20]. Since the Juelich atlas does not cover the whole brain, we focused on well-established areas, namely the primary somatosensory cortex including area 1, 2 and 3a and 3b; premotor cortex including area 6; visual cortex area including V1, V2, V3v, V4; and V5 and BA 44/45 area (Broca’s area).

The areas in cytoarchitectonic atlas are bordered in black dotted lines, and parcels having at least 70% overlap with cytoarchitectonic areas are in solid colors. The parcel boundaries align well with the primary somatosensory area, though part of the parcels extended into the posterior boundaries of the postcentral gyri,

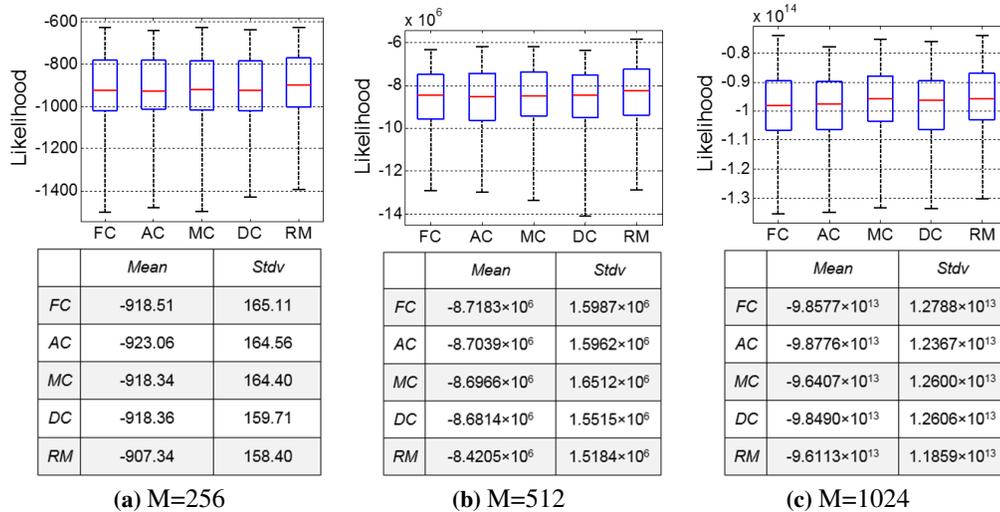
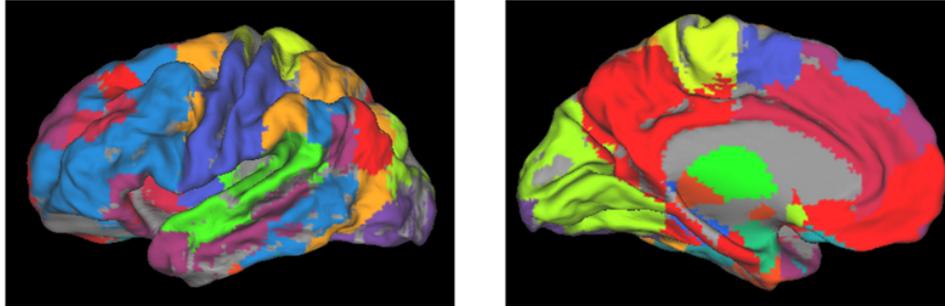
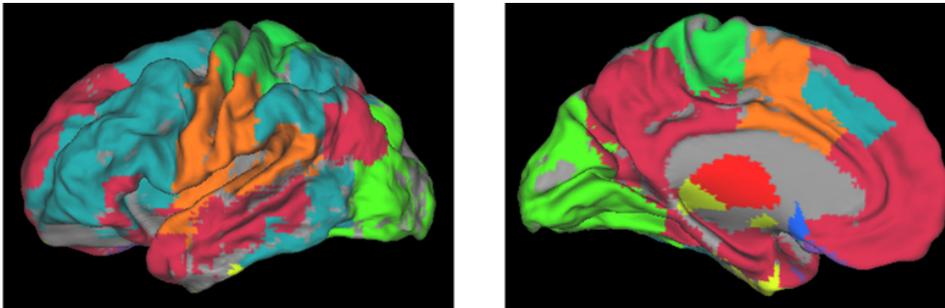


Figure 2.15: Leftout log-data likelihood comparison between multimodal and unimodal parcellation. Number of parcels was set to 256, 512 and 1024, respectively. (a) The likelihood of MC was higher than FC and DC, and significantly higher than AC at the scale of 256 parcels. (b) The likelihood of MC was higher than FC and AC at the scale of 512 parcels. (c) The likelihood of MC was significantly higher than FC, AC and DC at the scale of 1024 parcels. RM performs better due to the regular sampling of volume of interest from parcels with roughly equal size.

Figure 2.16 a. The Broca's area is largely covered by the parcels, with part of the parcels extending into the frontal cortex, Figure 2.16 b. The parcels also conformed well to the histological areas for the majority of the visual cortex, especially the primary visual cortex across the central visual field, Figure 2.16 c. For outer visual cortex, the parcels slightly extend beyond the histological areas. The parcels well cover the histological areas of sensorimotor areas with each parcel approximately corresponding to a functional subdivision of the motor map that represents a different part of the human body, the homunculus [173], Figure 2.16 d.



(a) Parcels grouped into 25 subnetworks



(b) Parcels grouped into 9 subnetworks

Figure 2.17: Extracted subnetworks. Subnetworks, such as Visual Network (light green in Occipital Cortex), Ventral Motor and Sensory system and Auditory system (orange), Dorsal Motor and Sensory system (spring green), and the Default Mode network (dark red) have been found in (b).

of less well-known systems (light purple) implicated in memory retrieval that have been detected in more recent studies [18, 69]. Furthermore, we tested the consistency of subnetwork structure generated from the two exclusive datasets from HCP. The subnetwork overlap was 80.26%. Note that based on maximal modularity, the “optimal” number of subnetwork was found to be 9 for the datasets used, Figure 2.17 b.

2.2.9 Discussion

2.2.9.1 Purposes and Key Challenges of Multimodal Parcellation

Through evolution, the human brain has been structured in a certain way for executing various functions. The cortical foldings, the white matter wiring, and the electrical and chemical signaling mechanisms all pertain to this complex organization. Thus, jointly exploiting all these brain attributes should theoretically generate parcellations that better resemble the brain's inherent division as compared to examining each attribute in isolation. One of the key challenges to adopting a multimodal approach is drawing the "right" balance between different brain attributes. We proposed basing this weighting on test-retest reliability of each attribute, which we empirically showed to work well in practice. More knowledge of the relationship between different attributes warrants further research. Another key challenge, which applies to brain parcellation in general, is the choice on the number of parcels. Choosing a criterion for finding the "ground truth" number of parcels is nontrivial. In fact, whether such a number exist is debatable, i.e., this choice is largely a matter of the spatial resolution that is suitable for one's analysis. Our take on this issue is not to settle on any particular number, but to adopt a multi-scale strategy. Findings that are consistent across scales are less likely to be false positives. We observed that the reproducibility Dice coefficients decreased as the number of parcels increase (from 256, 512 to 1024), since the combinatorial increase in possible voxel grouping decreases the probability of accurate assignments.

2.2.9.2 Multimodal Parcellation Improves Reproducibility and Data Likelihood, and Maintains Functional Homogeneity

Our results show that complementing functional connectivity with anatomical connectivity and structural information improves the reproducibility of the parcellations. Our reproducibility analysis also addresses the important question of parcellation stability with respect to sample perturbations. The high inter-subject test-retest reliability obtained shows that our multimodal scheme produces consistent parcellations for different subject subsets. Although multimodal informa-

tion would presumably disrupt maximization of functional homogeneity, the functional homogeneity of our multimodal-based parcellations was on par with pure functional parcellations, and higher than those based on anatomical connectivity, physical distance, randomly generated, and most of the existing atlases. The only exception is Gordon’s parcellation [163], but its higher functional homogeneity is likely due to the bias arising from parcellating only a thin shell of the cortex (see Figure 2.13). Further, the multimodal-based parcellations achieved higher leftout data likelihood than parcellations based on AC and surprisingly FC as well even though this evaluation metric is solely based on functional data.

2.2.9.3 Linking Parcels to Prior Knowledge

The subnetworks extracted from our multimodal-based parcellations match well with a good number of established functional systems [18, 69], which enabled us to assign neuroanatomically-meaningful labels to the parcels. Currently, we used only the functional data for subnetwork extraction. An important next step (presented in Chapter 5) would be to extend our multimodal scheme in grouping the parcels into subnetworks, and compare that with pure functional subnetwork extraction as well as existing multimodal methods. Multimodal parcel boundaries also align with those of cytoarchitectonic areas, despite that we did not explicitly optimize for this overlap in multimodal parcellation scheme. We suspect this overlap arose from the correlation between cytoarchitecture and brain function [174]. As for the misalignments, a likely reason is inter-subject misregistration. Take the BA 44/45 area for example, the misregistration errors are in the order of 6–12 mm reported in previous studies [61]. Also, there exist transition regions between cytoarchitecture areas [61] (not display in Figure 2.16). We further note that each cytoarchitectonic area is covered by multiple parcels, as opposed to a single parcel, which is likely due to functional subdivisions within each cytoarchitectonic area that are not apparent from the cytoarchitectonic attributes [163].

2.2.9.4 Region Level vs. Whole-brain Parcellation

We presented a region level extension to our multimodal parcellation approach, which greatly reduces computational cost and provides the side-benefit of incorpo-

rating structural information on gyri and sulci. This extension, which uses whole-brain connectivity to separately parcellate each HO region, intrinsically boosts reproducibility since both the number of voxels and the number of parcels are substantially reduced, which decreases the possible grouping arrangements, simplifying the problem of parcel correspondence between subjects. Also, part of the boundaries for some parcels would necessarily be confined to the boundaries of the predefined regions. To ensure fair comparisons, we applied exactly the same region level strategy for the contrasted unimodal parcellation methods and showed that multimodal information indeed provides substantial gains. Also, this region level strategy reduced computational time from 30 minutes for simultaneous whole-brain parcellation to 10 minutes with a 64-bit Windows machine on an Intel Xeon E3-1240V2 3.40GHz processor with 32GB of RAM. The main caveat is that brain units spanning multiple predefined structural brain regions would be split.

2.3 Summary

Towards improving the brain node definition, i.e., parcellation, we strive to tackle the challenges mentioned in Section 1.5 from two aspects.

In order to devise a more reliable method for parcellation, we embedded neighborhood connectivity information into the affinity matrix to ameliorate the adverse effects of noise. Our proposed approach produces parcellations with better intra-subject test-retest reliability, higher inter-subject test-retest reliability and higher subject-group consistency.

We investigated the important question of whether combining modalities can improve brain parcellation. We proposed an approach for integrating anatomical and functional connectivity information for brain parcellation. Applying our multimodal approach to the HCP data, we quantitatively showed the superiority of multimodal parcellation compared to unimodal parcellation and existing atlases in terms of reproducibility, functional homogeneity and leftout data likelihood. Multimodal parcellations also overlap with known cytoarchitectonic areas and the extracted subnetworks matched well with the established brain systems. Collectively, our results demonstrated that integrating multiple brain attributes that intrinsically reflect the brain's organization does indeed improve brain parcellation. Our mul-

timodal volumetric parcellations could thus serve as excellent complements to the recent cortical multimodal surface parcellations, especially considering the importance of subcortical regions for clinical investigations, such as Parkinson's disease studies.

Chapter 3

Noise Reduction for Brain Network Edge Building

This chapter focuses on improving brain edge building, which is based on paper [P4] and the first part of paper [P5]. The connectivity estimates depend greatly on the quality of the fMRI or dMRI data, both of which suffer from pronounced noise and image resolution limitations. To tackle the problems in reducing the false negatives and positives in connectivity estimates, we propose a matrix completion based technique to combat false negatives by recovering missing connections, and present a local thresholding method which can address the regional bias issue when suppressing false positives in connectivity estimates.

3.1 Matrix Completion to Combat False Negatives

3.1.1 Related Work

Due to the limited spatial resolutions, complexities of the underlying tissue and uncertainties of signal noise [83] (detailed in paragraph 1.4.2.3.1), Anatomical Connectivity estimates derived from the fiber tract strength suffer from both false positive and negative connection problem, especially the false negatives. In particular, diffusion direction is ambiguous at the crossing fiber locations, causing pre-mature termination of tracts with conventional streamline algorithms [13]. To better handle

crossing fibers, techniques for improving ODF estimation [175] and tractography [74, 79] have been proposed. However, even if the crossing fiber issue is resolved, the ambiguity in diffusion direction near the gray-white matter interface introduces great uncertainty to the fiber endpoint locations [13]. Heuristics for endpoint extrapolation have been explored with only modest improvement shown [150]. By and large, due to crossing fibers and fiber endpoint uncertainty, AC estimates tend to be more prone to missing connections, i.e., false negatives [13].

A fundamental limitation to the above strategies for improving AC estimation is the upper bound inherently imposed by attainable imaging resolutions. Endeavors for increasing the image spatial resolution include improving the acquisition scheme [146, 176] and applying super-resolution techniques [177]. However, due to the physical limitations of MRI technology, micro-meter resolution, i.e., the width of a fiber, is currently unattainable for whole-brain coverage. Even if such fine resolution is attained, the required computational time for tractography would be impractical.

3.1.2 Low Rank Matrix Completion for Connectivity Estimation

We propose here a matrix completion based approach for recovering missing connections. The underlying assumption is that the missing connections are intrinsically related to the observed entries of the AC matrix, hence can be recovered using matrix completion.

In matrix completion problems, one attempts to fill missing entries of a partially observed matrix. The problem is underdetermined if one does not restrict the degrees of freedom. The typical way to impose such restrictions is by minimizing the rank:

$$\min_{\mathbf{X} \in \mathbb{R}^{m \times n}} \text{rank}(\mathbf{X}), \text{ s.t. } \mathbf{X}_{i,j} = \mathbf{M}_{i,j}, \forall (i,j) \in \Omega \quad (3.1)$$

while constraining the matrix entries, $\mathbf{X}_{i,j}$, to match the observed values, $\mathbf{M}_{i,j}$. One has available s sampled entries $\{\mathbf{M}_{i,j} : i,j \in \Omega\}$ where Ω is a random subset of cardinality s . Since Equation 3.1 is Non-deterministic Polynomial-time (NP) hard, a convex relaxation is often employed [178]:

$$\min_{\mathbf{X} \in \mathbb{R}^{m \times n}} \|\mathbf{X}\|_*, \text{ s.t. } \mathbf{X}_{i,j} = \mathbf{M}_{i,j}, \forall (i,j) \in \Omega \quad (3.2)$$

where $\text{rank}(\mathbf{X})$ is approximated by the nuclear norm (also known as the trace norm), $\|\mathbf{X}\|_*$, i.e., sum of the singular values of \mathbf{X} . One strategy for solving Equation 3.2 that well suits the AC estimation problem is to find a low rank matrix, $\mathbf{X} = \mathbf{Y}\mathbf{Z}$, that minimizes $\|\mathbf{X} - \mathbf{M}\|_F^2$ [179], where $\mathbf{Y} \in \mathbb{R}^{m \times r}$, $r < m$, and $\mathbf{Z} \in \mathbb{R}^{r \times n}$, $r < n$. In the present context where $\mathbf{M} \in \mathbb{R}^{N \times N}$ is an AC matrix of N brain regions, \mathbf{X} should factorize as $\mathbf{Y}\mathbf{Y}^\top$ given the symmetry of the AC matrix, where each column i of $\mathbf{Y} \in \mathbb{R}^{N \times r}$ can be interpreted as the membership weights of N brain regions belonging to subnetwork i . The rank, r , thus corresponds to the number of subnetworks identifiable from \mathbf{M} . The detailed optimization process can be found in [179]. To ensure symmetry, $(\mathbf{X} + \mathbf{X}^\top)/2$ is taken as our recovered AC estimates. We refine all AC connections by updating their values with the matrix completion corrected values. The factorization algorithm [179] we used for minimizing the matrix completion objective was only one of many possible optimization schemes. This algorithm empirically gives us the best performance of missing entry recovery based on our experiments in Section 3.1.5.2.

3.1.3 Rank Range Search and Aggregation

Intuitively, AC matrix rank corresponds to the number of brain subnetworks. Ample studies suggested that the brain comprises only about a dozen large-scale subnetworks, which confirms the low rank assumption. Existing rank selection methods are largely based on finding a transition point in the matrix eigenspectrum [180]. However, AC matrices typically do not display clear transition points (Figure 3.1). To mitigate poor recovery due to choosing the “wrong” rank, we propose here to aggregate recovered entries over a range of ranks. The intuition is that since matrix rank corresponds to the number of subnetworks, which arguably is a matter of resolution, aggregating over ranks in effect integrates modularity information across different subnetwork refinements. To select a rank range, we present the following automated strategy. Over a predefined range, we first find the rank that provides the best recovery accuracy, a , based on removing a percentage of the observed entries and assessing how well those entries are recovered. The range is then defined as all ranks $r \in [r_1, \dots, r_k]$ with accuracy within $[a - \varepsilon, a]$. We then aggregate the recovered AC matrices for this rank range by taking their median

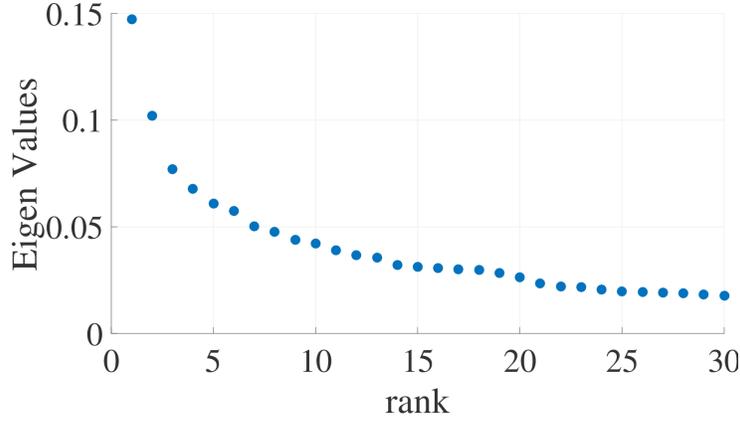


Figure 3.1: Eigenspectrum of an exemplar AC

value: $\bar{\mathbf{X}}_{i,j} = \text{median}(\mathbf{X}_{i,j}^r, r \in [r_1, \dots, r_k])$, where $\mathbf{X}_{i,j}^r$ is the estimated $\mathbf{X}_{i,j}$ for rank r .

3.1.4 Negative Entries Filling using Neighborhood Information

Standard matrix completion algorithms do not impose a non-negative constraint on matrix entries [181]. However, as for the AC estimation, negative recovered entries are not biologically interpretable, since fiber counts cannot be negative. Explicitly imposing a non-negative constraint onto matrix completion is possible [181], but we observe empirically that such constraint tends to decrease the recovery accuracy (Figure 3.2). The possible explanation could be the fiber count properties do not fit well in a non-negative low rank matrix completion setting. Instead, we devise a method to exploit those few entries which are estimated as negative based on neighborhood information. Under the hypothesis that a pair of highly interlinked brain regions are connected to similar brain areas, we first compute a Pearson’s correlation matrix from $\bar{\mathbf{X}}$: $\mathbf{C} = \bar{\mathbf{X}}\bar{\mathbf{X}}^\top$, for defining neighbors based on similarity in connection patterns $\bar{\mathbf{X}}_{i,:}$. For each negative recovered entry, $\mathbf{X}_{i,j}$, we search for h regions with the highest positive correlation to region j , $H_j = \{j_1, j_2, \dots, j_h\}$, where h is selected using cross-validation with recovery accuracy as the metric. The negative recovered entries are then interpolated by taking a weighted mean of

h positive recovered entries between region i and region j 's neighbors, H_j :

$$\hat{\mathbf{X}}_{i,j} = \begin{cases} \frac{\sum_{t=1}^h (\bar{\mathbf{X}}_{i,jt} \mathbf{C}_{i,jt})}{\sum_{t=1}^h (\mathbf{C}_{i,jt})}, & \text{if } \bar{\mathbf{X}}_{i,j} < 0 \\ \bar{\mathbf{X}}_{i,j}, & \text{otherwise} \end{cases} \quad (3.3)$$

3.1.5 Experiments

We refer to our approach as MCmedFill (Matrix Completion with median aggregation over candidate ranks and neighborhood information filling), which we validate on synthetic data using recovery accuracy, and apply to real data from the HCP [146] for IQ prediction.

3.1.5.1 Materials

3.1.5.1.1 Synthetic Data We generated 100 synthetic datasets that cover a variety of network configurations. Each network comprised $N=112$ regions analogous to the HO atlas [57]. For each dataset, we set the number of subnetworks, M , to a random value between 12 and 14, conforming to current literature [23]. The number of regions in each subnetwork was set to $(N/M) + q$ with q being a random number between $[-2, 2]$. With the resulting configuration, we created the corresponding adjacency matrix, Σ , and added Gaussian noise, Figure 3.2a. Negative matrix entries were set to zero. Lastly, we randomly set 20% of the ground truth connections to 0 to model how AC estimates are prone to false negatives, Figure 3.2b.

3.1.5.1.2 Real Data We used the dMRI scans and fluid IQ scores of 77 healthy subjects from the HCP Q3 dataset [146], see details in Section 1.8. Given the pre-processed dMRI data [146], we applied global tractography based on CSA ODF and Gibbs tracking [81] using the MITK package [149]. To define brain regions, we employed the HO atlas [57], which has 112 regions. To compute the fiber count between regions, we warped the HO atlas onto the $b = 0$ volume of each subject using affine registration. We further extrapolated the fiber endpoints using Gaussian

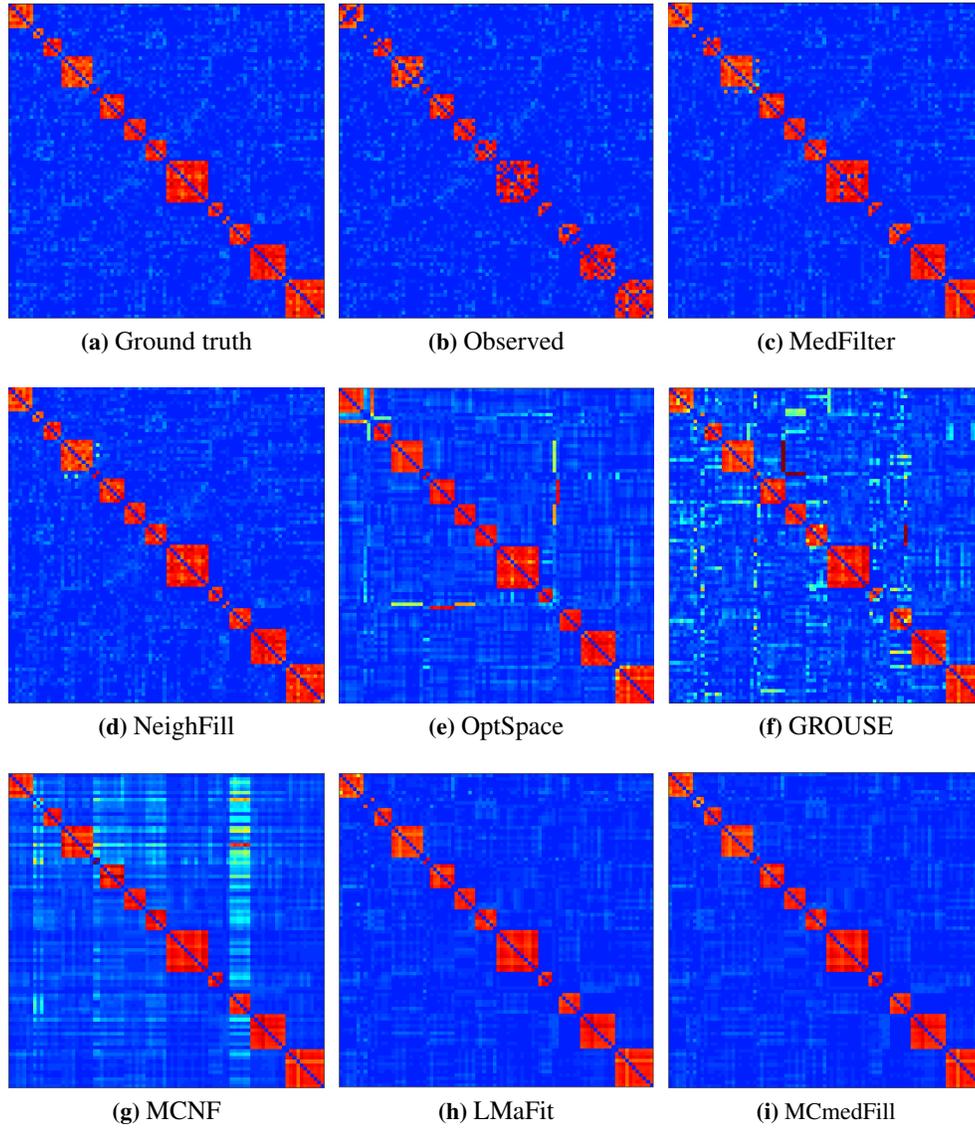


Figure 3.2: Matrix recovery on an exemplar synthetic dataset. MCmedFill (i) more accurately recovered the ground truth connections (in red) than the contrasted methods (c-h).

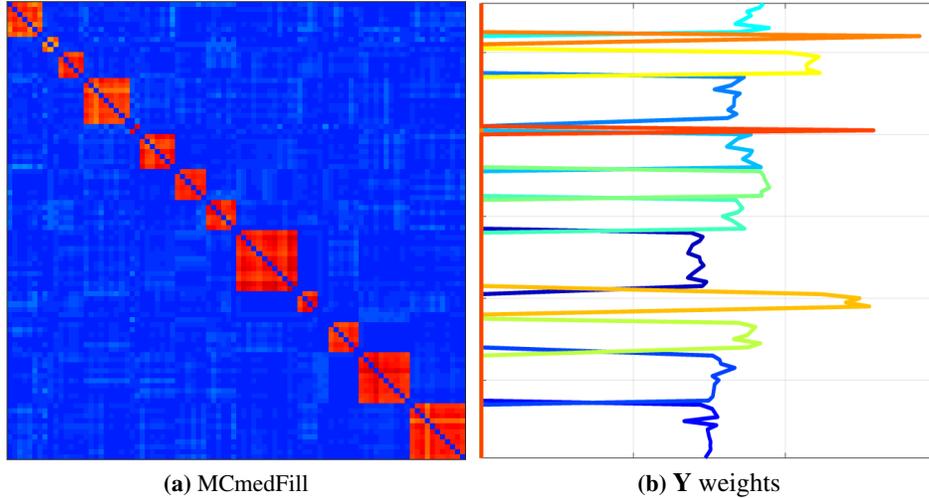


Figure 3.3: Matrix recovery using MCmedFill (a) on an exemplar synthetic dataset. The weights in each column of the low rank matrix \mathbf{Y} align with the ground truth subnetwork assignment. (b) An example of thresholded weights for MCmedFill shown.

kernels [150]. The resulting fiber counts normalized by the region volume were taken as our AC estimates.

3.1.5.2 Results

To execute our approach, MCmedFill, we used Low-Rank Matrix Fitting (LMaFit) [179] for matrix completion due to the subnetwork analogy that it provides (Section 3.1.2) as well as its computational speed. Empirically, we set $\varepsilon = 0.2a$ and predefined rank range as [1, 30]. We first evaluated MCmedFill based on recovery accuracy on synthetic and real data. Since brain connectivity presumably relates to IQ [182], we further evaluated MCmedFill based on IQ prediction on real data. For comparison, we examined interpolating AC entries with zero values by median filtering (MedFilter), with neighborhood information as described in Section 3.1.4 (NeighFill), and three other widely-used rank-based matrix completion algorithms: OptSpace [180], Grassmannian Rank-One Update Subspace Estimation (GROUSE) [183], and Matrix Completion with Nonnegative Factoriza-

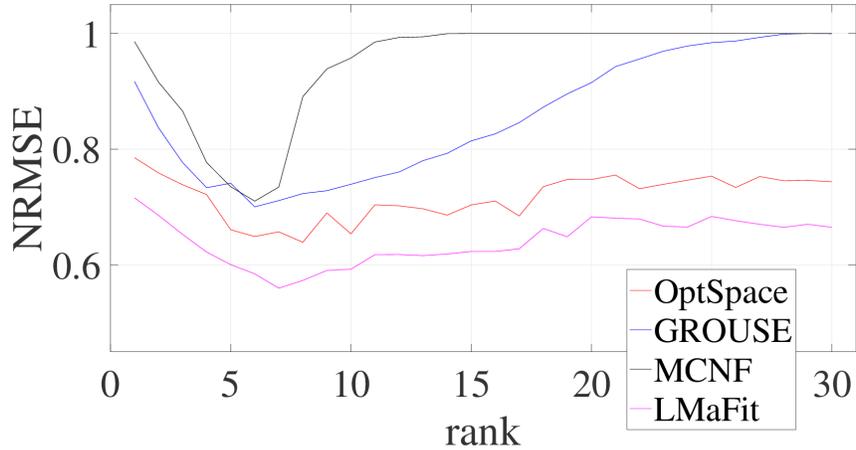


Figure 3.4: Rank selection method. Our criteria of using recovery accuracy on randomly removed AC entries displays clear NRMSE minima for rank selection.

tion (MCNF) [181], which imposes a non-negative constraint. The vanilla LMaFit was also tested. The rank for OptSpace, GROUSE, MCNF, and vanilla LMaFit were selected based on highest recovery accuracy over a range of ranks (Figure 3.4). All statistical comparisons were based on the Wilcoxon signed rank test with MCmedFill as the reference. Significance was declared at an α of 0.05 with Bonferroni correction.

3.1.5.2.1 Recovery Accuracy We assessed recovery by first deleting 20% of the entries in AC matrices for both synthetic and real data. We then estimated the recovery accuracy using: normalized root-mean-squared-error (NRMSE).

$$\text{NRMSE} = \sqrt{\sum(\mathbf{M}_{i,j} - \hat{\mathbf{X}}_{i,j})^2} / \sqrt{\sum \mathbf{M}_{i,j}^2}$$
, where \mathbf{M} is the ground truth and $\hat{\mathbf{X}}$ is the recovered matrix, and coefficient of determination: $R^2 = 1 - SS_{res}/SS_{tot}$, where $SS_{res} = \sum(\mathbf{M}_{i,j} - \hat{\mathbf{X}}_{i,j})^2$ and $SS_{tot} = \sum(\mathbf{M}_{i,j} - \bar{\mathbf{M}}_{i,j})^2$. Lower NRMSE and higher R^2 indicate higher accuracy.

On the 100 synthetic datasets, MCmedFill achieved significantly higher accuracy than the contrasted methods, Figure 3.5a, 3.5b. For real data, we first generated a group AC matrix by averaging the subject level AC estimates. We then randomly deleted 20% of the entries for 100 times. MCmedFill again achieved

significantly higher accuracy, Figure 3.6a, 3.6b. Note the reported accuracies for the four matrix completion algorithms were the best results across a range of ranks (Figure 3.4). Also, MCmedFill’s improvement over LMaFit might appear small, but statistically significant since higher accuracies were consistently observed across the random test cases.

Examining the columns of the low rank matrix \mathbf{Y} , we observed that the weights (thresholded at 50% of the maximum value) match the ground truth of subnetwork assignment for synthetic data (Figure 3.3b) using optimally selected rank, which happens to correspond to the ground truth number of subnetworks. Furthermore, some columns of \mathbf{Y} were found to resemble known brain networks, e.g., Frontal-Parietal network (Figure 3.7b), suggesting the potential of using matrix completion for subnetwork extraction, which we defer for future work.

3.1.5.2.2 IQ Prediction Due to the absence of a ground truth human anatomical connectivity atlas, we further assessed MCmedFill based on IQ prediction with real data. We used the lower triangular entries of the recovered subject level AC matrices as predictors. Confounds including age and sex were regressed out from the predictors and the IQ scores prior to applying L2-regularized L2-loss Support Vector Regression (SVR) [184]. SVR solves the following primal problems:

$$\min_w 1/2w^T w + C \sum_{i=1}^l (\max(0, |y_i - w^T x_i| - \varepsilon))^2, \quad (3.4)$$

where y_i is the real value of the IQ score, $x_i \in \mathbb{R}^n$, $C > 0$ is a penalty parameter and ε is a parameter to specify the sensitiveness of the loss. 100 random realizations of 11-fold cross-validation were performed for estimating prediction accuracy and its variability. Prediction accuracy was defined as the Pearson’s correlation, R , between the 77 predicted and observed IQ scores for each cross-validation realization. Though indirect, we are very thrilled to show that our validation approach resulted in significant findings on real data that are neuroscientifically meaningful.

Predicting IQ with the original AC estimates obtained an average prediction correlation of 0.1865, Figure 3.7a. Applying MedFilter, OptSpace, and GROUSE degraded the prediction. NeighborFill and LMaFit performed slightly better than

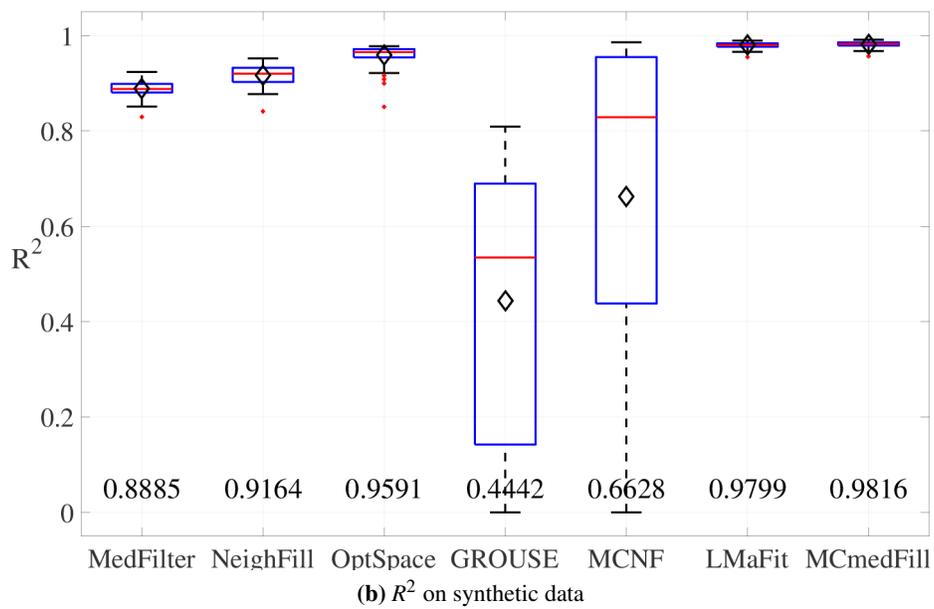
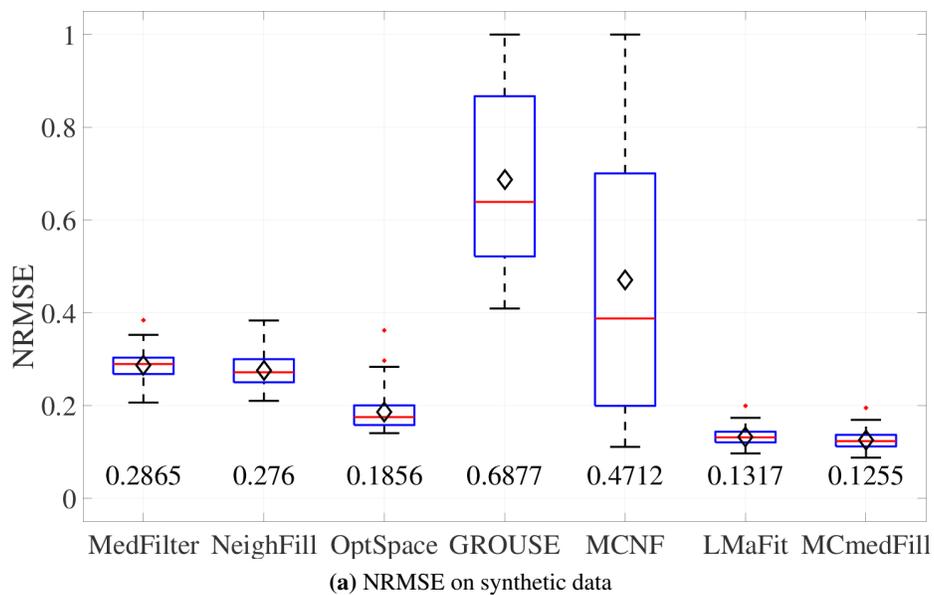
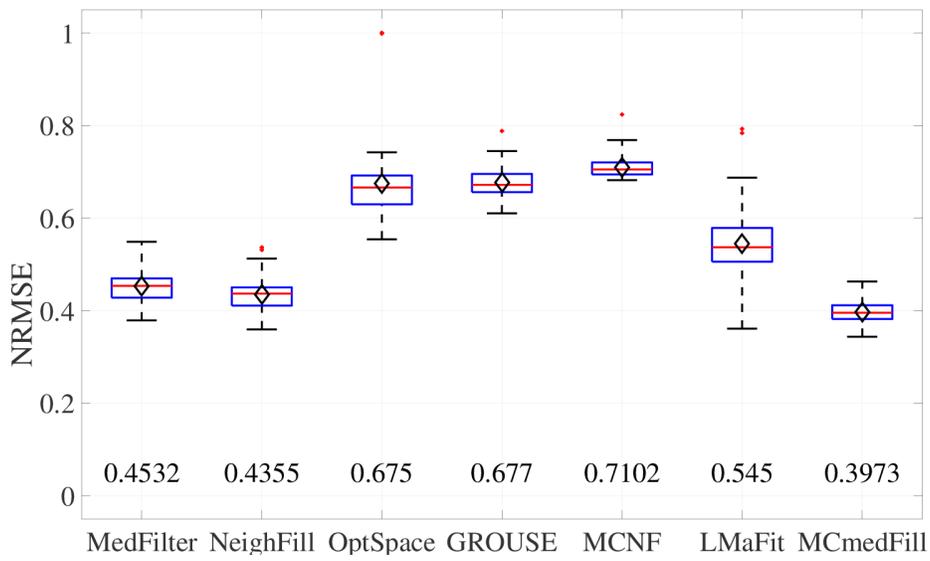
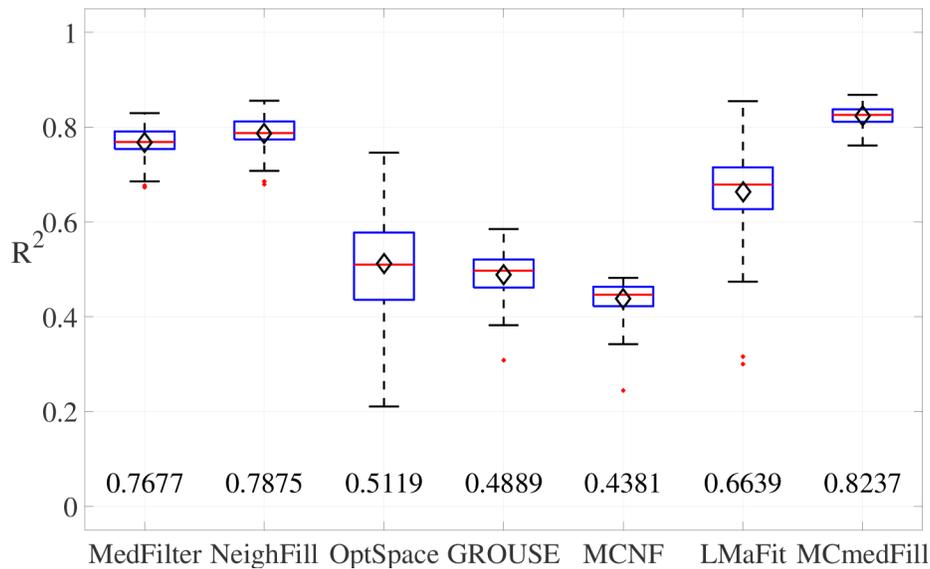


Figure 3.5: Matrix recovery on synthetic data. Mean values are indicated as black diamonds and labeled at the bottom of each subfigure. MCmedFill achieved significantly higher accuracy than contrasted methods based on Wilcoxon signed rank test.

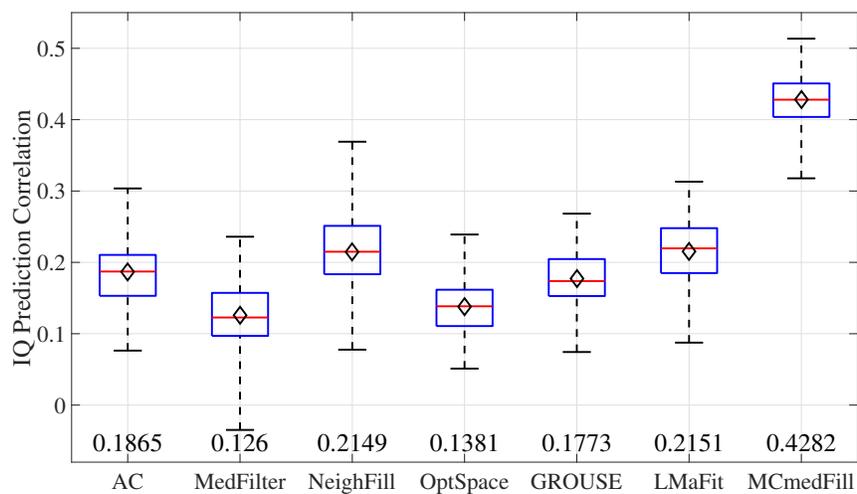


(a) NRMSE on real data from HCP

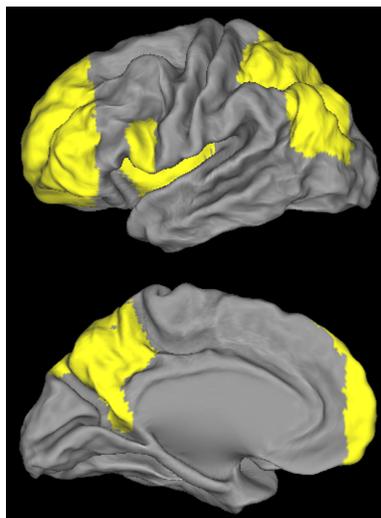


(b) R^2 on real data from HCP

Figure 3.6: Matrix recovery on real data. Mean values are indicated as black diamonds and labeled at the bottom of each subfigure. MCmedFill achieved significantly higher accuracy than contrasted methods based on Wilcoxon signed rank test.



(a) IQ prediction on real data from HCP



(b) \mathbf{Y} weights visualization

Figure 3.7: IQ prediction on real data. (a) MCmedFill achieved significantly higher correlation coefficient between observed and predicted IQ scores than contrasted methods. (b) Brain region subnetwork weights along columns of \mathbf{Y} (thresholded for clearer visualization) were found to resemble known brain networks, e.g., Frontal-Parietal network.

original AC estimates, but not by a statistically significant amount. In contrast, MCmedFill achieved significantly higher prediction accuracy than the contrasted methods, with an average prediction correlation of 0.4282 attained, which is higher than values reported in most AC-based IQ prediction studies [182].

To investigate the sources of prediction improvement, we contrasted the AC matrices before and after applying MCmedFill with a focus on connections between regions within the Default Mode Network (DMN) as well as within the Executive Control Network (ECN) (both networks had been shown to relate to IQ [182]). Applying MCmedFill resulted in a 53.19% and a 28.41% increase in estimated connectivity within DMN and ECN, respectively, which suggests that the improved IQ prediction with our approach might be due to the recovery of relevant connections that were missed with standard AC estimation.

3.1.6 Discussion

We proposed a matrix completion based approach for recovering missing connections to improve the AC estimation. By aggregating recovered entries over ranks and interpolating negative entries with neighborhood information, MCmedFill attained higher accuracy in recovering deleted AC entries for both synthetic and real data. Higher accuracy in IQ prediction was also shown. Our results thus demonstrated clear benefits of refining conventional AC estimates with our approach.

The study on the capability of our approach to deal with the false positives that also exist in AC estimates [185] needs to be continued. The low rank assumption has a denoising effect, which could reduce the effect of false positives. On the other side, we need to explore whether the proposed method could negatively affect the false discovery rate, which can impact the clinical feasibility of using anatomical connections to study some real life diseases. We will also work on more systematic and detailed validation, such as direct comparison between publicly available macaque dMRI connections and tract tracing results [186, 187]. Tract tracing provides a gold standard of brain connections, so the accuracy of our approach can be evaluated directly on macaque data.

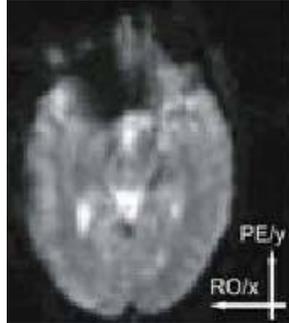


Figure 3.8: Signal dropouts in the orbitofrontal cortex (OFC) caused by field inhomogeneities (at 1.5 T). Image courtesy of [21].

3.2 Local Thresholding to Suppress False Positives

3.2.1 Related Work

The human brain regions and their pair-wise interactions constitute graph nodes and weighted edges, respectively. fMRI is one of the widely used modalities for connectivity estimation, however, FC estimation is known to suffer especially from false positives [90]. Moreover, confounds, such as region size bias [188, 189], effects of motion artifacts [190], and signal dropouts due to susceptibility artifacts (especially in regions like the orbitofrontal cortex, Figure 3.8, and the inferior temporal lobe) [21], introduce region-specific biases to the connectivity estimates. Such problems make applications such as subnetwork extraction very challenging, since brain network topology may be obscured by noisy connectivity estimates [16].

The conventional way for dealing with noisy connectivity matrices is to apply Global Thresholding (GT) by either keeping only connections with values above a certain threshold or keeping a certain graph density [16]. Due to region-specific connectivity biases, e.g., brain regions in signal dropout locations tend to display lower connectivity, certain regions that do belong to a subnetwork might not appear as such based on the fMRI measurements, especially after GT, which prunes weak edges. To mitigate this overlooked problem, a Local Thresholding (LT) method based on the MST-KNN has been proposed [22]. The idea in [22] was to build a

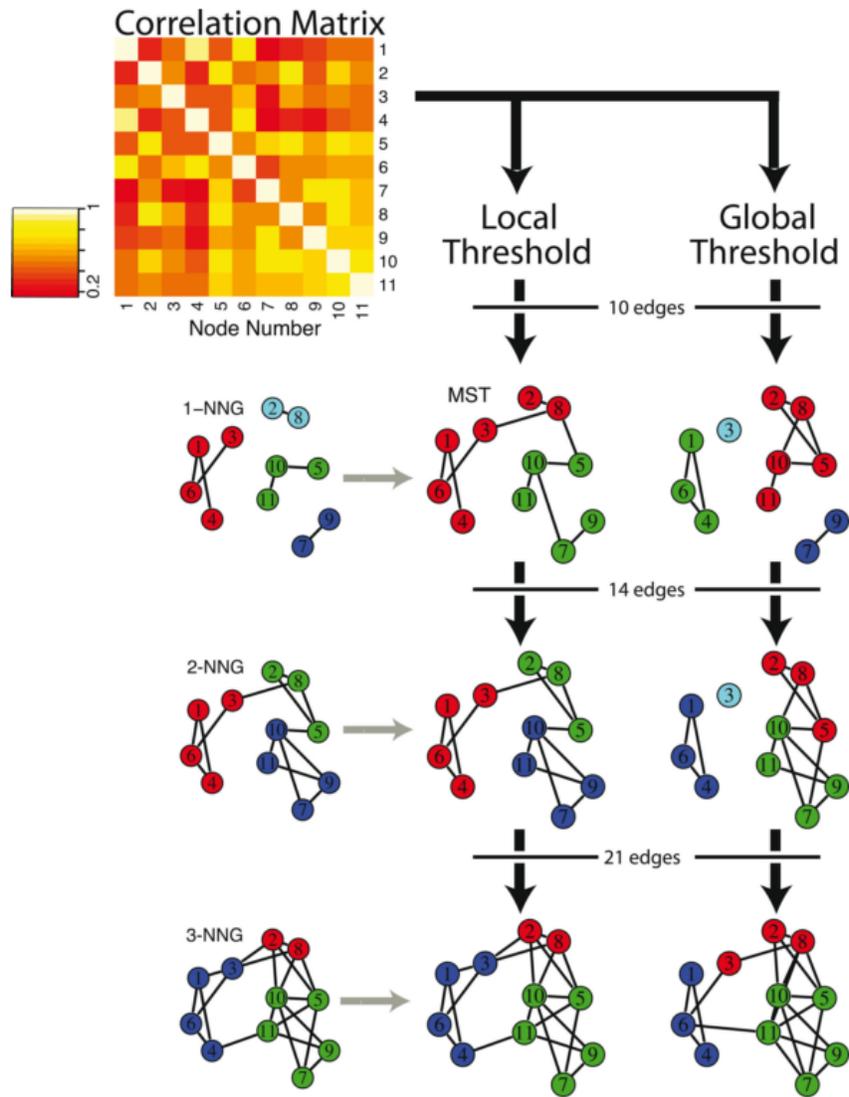


Figure 3.9: Schematic illustrating local thresholding (MST-KNN) and global thresholding methods. Image courtesy of [22].

single connected graph using the minimum spanning tree and expand the tree by adding edges from each node to its nearest neighbors until a desired graph density is reached, Figure 3.9. However, the key step of adding edges to all nodes when expanding the tree, based on the assumption of equal density/importance for each node, lacks neuroscientific justifications. A few studies have explored spectral graph wavelet transform for graph denoising [191], but this approach does not explicitly handle region-specific connectivity biases. In fact, most existing connectivity estimation [16] do not account for these biases.

3.2.2 Local Thresholding

To deal with noisy edges mostly with false positives and region-specific connectivity biases, we propose a local thresholding scheme that normalizes the connectivity distribution of each node prior to thresholding. Due to region-specific connectivity biases, conventional GT might prune relevant connections with weak edge strength. To account for these biases, we present here a LT scheme. The idea is to first normalize the connectivity distribution of each node into a uniform interval to rectify the biases. Subsequent Global Thresholding on this normalized graph would have the effect of applying Local Thresholding on each node without enforcing the equal importance of all nodes. Specifically, let \mathbf{C} be an $N \times N$ connectivity matrix, where N is the number of nodes in the brain graph. We normalize the connectivity distribution by mapping each row of \mathbf{C} from $[\min(\mathbf{C}_{i,:}), \max(\mathbf{C}_{i,:})]$ to $[0, 1]$, where $\mathbf{C}_{i,:}$ denotes row i of \mathbf{C} corresponding to the connectivity between brain region i and all other regions in the brain. A threshold is then applied to generate a binary adjacency matrix, \mathbf{G} , which we then symmetrize by taking the union of \mathbf{G} and \mathbf{G}^T : $\mathbf{A} = \mathbf{G}_{i,j} \cup \mathbf{G}_{j,i}$, where \mathbf{G}^T is the transpose of \mathbf{G} . This binary adjacency matrix \mathbf{A} is used to mask out the noisy edges from \mathbf{C} : $\hat{\mathbf{C}}_{i,j} = \mathbf{A}_{i,j} \mathbf{C}_{i,j}$, which is equivalent to applying a local threshold to $\mathbf{C}_{i,:}$ for all i . We note that in the event that noisy nodes (signal dropout brain regions, Figure 3.8) are accidentally included, some of the connections to these noise nodes (that might not be kept by GT) would be kept by LT due to the normalization step.

3.2.3 Experiments

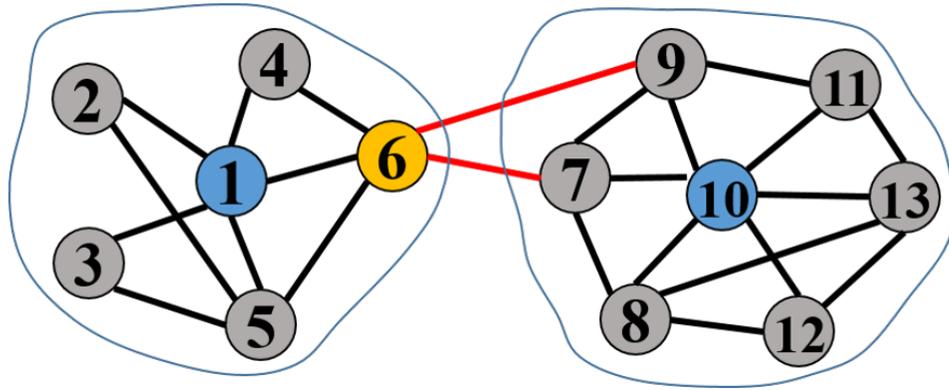
We validated our LT scheme using subnetwork extraction application based on the assumption that we can achieve more accurate subnetwork structure using more accurate connectivity estimates.

3.2.3.1 Materials

3.2.3.1.1 Synthetic Data To illustrate our strategy, we synthesized a small-scale network consisting of $N = 13$ nodes in Figure 3.10 a. These nodes are assigned to two subnetworks with each subnetwork having a provincial hub (blue) and linked by a connector hub (orange). The weights of edges in the network were set to 0.75. We then added Gaussian noise at a SNR of -5 dB to simulate the motion artifacts [190]. We further lowered the edge weights of node 2 and 12 by 20% to simulate region-specific connectivity biases for smaller brain regions [188], Figure 3.10 b.

We also generated synthetic data that cover 100 random network configurations with N set to 100 nodes. For each network configuration, the number of subnetworks, M , was randomly selected from [10, 20]. The number of regions within each subnetwork was set to $\text{round}(N/M) + r_n$, where r_n was randomly selected from [-2, 2]. With the resulting configuration, we created the corresponding adjacency matrix, Σ , and drew time courses with 4,800 samples (analogous to real data) from $N(0, \Sigma)$. We then added Gaussian noise to the time courses with signal-to-noise ratio randomly set between [-6dB, -3dB]. Sample covariance was then estimated from these time courses with correlation values associated with $q\%$ of the nodes reduced by $z\%$, where q was randomly selected from [20%, 30%] and z was randomly selected from [30%, 40%] to simulate region-specific connectivity biases for smaller brain regions [188].

3.2.3.1.2 Real Data We used the rs-fcMRI data of 77 healthy HCP subjects, details and preprocessing done on the data described in Section 1.8. We used the Will90fROI atlas [23] to define ROIs, which have 90 ROIs being assigned to 14 well-established brain systems Figure 3.11. Voxel time courses within ROIs were averaged to generate region time courses. The region time courses were demeaned,



(a) Network structure

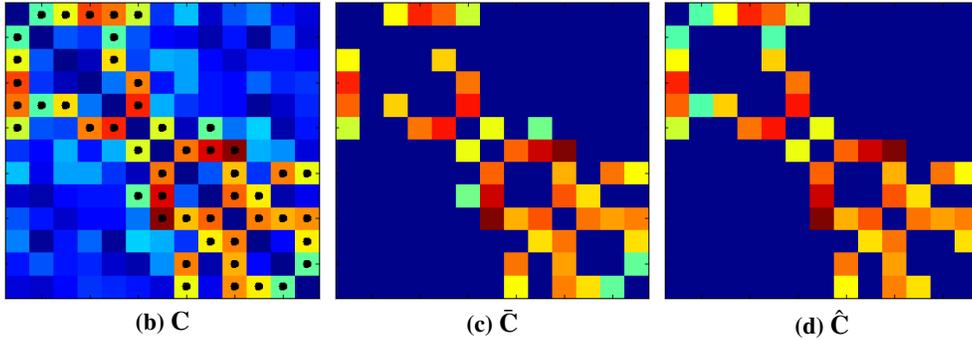


Figure 3.10: Schematic illustrating our proposed local thresholding using a small scale example having two subnetworks with each subnetwork having a provincial hub (blue) and linked by a connector hub (orange), shown in (a). In (b), warmer color indicates higher connectivity and black dots indicate the ground truth adjacency matrix. We denote \tilde{C} as global thresholded, and \hat{C} as local thresholded connectivity matrix. At a graph density of 0.25, the GT generated isolated node 2 in (c), while our LT preserved two edges linked to node 2 in (d).

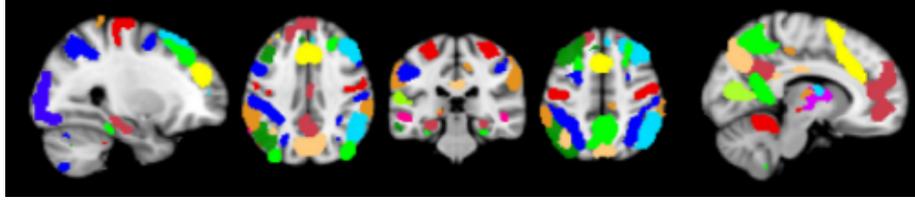


Figure 3.11: Will90fROI atlas [23] with 90 ROIs being assigned to 14 well-established brain systems.

normalized by the standard deviation, and concatenated across subjects for extracting group subnetworks. The Pearson’s correlation values between the region time courses were taken as estimates of connectivity. Negative elements in the connectivity matrix were set to zero due to the currently unclear interpretation of negative connectivity [93].

3.2.3.2 Results

We compared our LT strategy against GT and MST-KNN in [22]. Instead of using a specific threshold, we examine a range of graph densities to test the robustness of our proposed strategy. For synthetic data, evaluation was based on the accuracy of subnetwork extraction. To estimate accuracy, we matched the extracted subnetworks to the ground truth subnetworks using Hungarian clustering [162] with the $DSC = 2|X \cap Y| / (|X| + |Y|)$, where X is the set of regions of an extracted subnetwork and Y is the set of regions of a ground truth subnetwork. The average DSC over matched subnetworks was taken as accuracy.

For real data, we assessed the overlap between the extracted subnetworks and fourteen well-established brain systems [23] and subnetwork reproducibility for a range of graph densities on HO atlas [57] using DSC.

All statistical comparisons are based on the Wilcoxon signed rank test with significance declared at an α of 0.05 with Bonferroni correction.

3.2.3.2.1 Synthetic Data In the example of small-scale network, with GT (Figure 3.10 c), node 2 was isolated from subnetwork 1. In contrast, our LT scheme (Figure 3.10 d) was able to preserve node 2. Also, with our LT (Figure 3.10 d), one

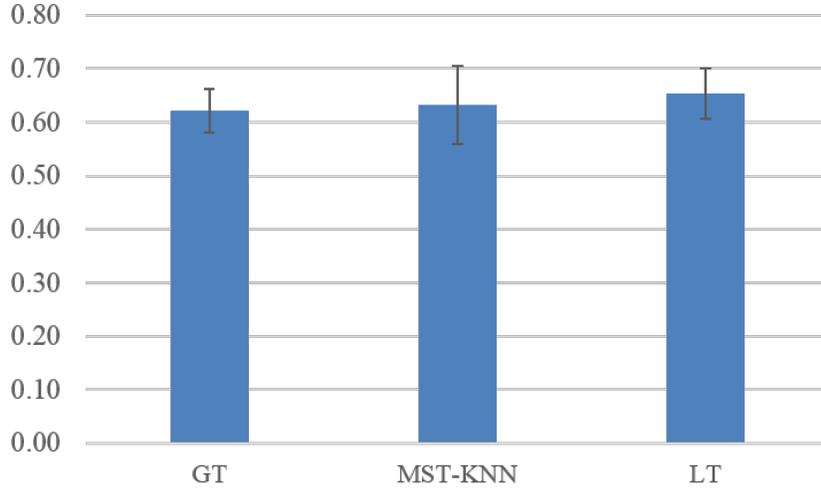
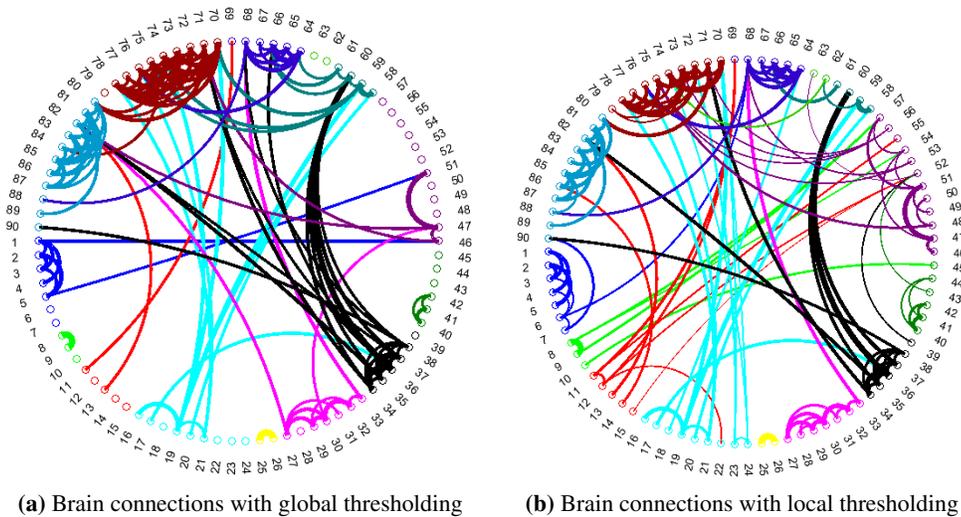


Figure 3.12: Subnetwork extraction accuracy using synthetic dataset using different thresholding schemes. The bar chart indicates the average DSC over 100 synthetic dataset across a graph density range of [0.005, 0.5] at an interval of 0.01.

of between-subnetwork edges (i.e., edges between nodes 6 and 7 & nodes 6 and 9) was pruned, which would help prevent the two subnetworks from being declared as one, whereas none of between-network edges was pruned using GT (Figure 3.10 c).

On the 100 synthetic dataset with 100 nodes over a density range of [0.005, 0.5] at an interval of 0.01, LT achieved statistically significantly higher accuracy (average DSC = 0.6537 ± 0.0479) than GT (average DSC = 0.6216 ± 0.0417), and MST-KNN (average DSC = 0.6327 ± 0.0732), Figure 3.12.

3.2.3.2.2 Real Data We first evaluated our strategy by examining the overlap between our extracted subnetworks and 14 well-established brain systems presented in [23], which we used as ground truth. Our LT achieved an average DSC of 0.5936 ± 0.0374 , which was statistically significantly higher than GT (average DSC = 0.5384 ± 0.0436), and MST-KNN (average DSC = 0.4567 ± 0.0653). We note that although some node-wise variations in connectivity distribution might have a neuronal basis, we postulate that these variations would be overwhelmed by the



Color	Name	Color	Name
blue	Anterior Salience	dark green	Sensorimotor
spring green	Auditory	purple	Posterior Salience
red	Basal Ganglia	dark cyan	Precuneus
cyan	dDMN	light green	Primary Visual
yellow	Higher Visual	navy	RECN
magenta	Language	dark red	vDMN
black	LECN	steelblue	Visuospatial

(c) Color labels of 14 well-established brain subnetwork

Figure 3.13: Circular plots of the whole brain connections based on the 14 subnetwork structure using different thresholding schemes. Exemplar results shown here are thresholded at the graph density of 0.02. The thickness of the connections indicate the strength of the connections.

various confound-induced connectivity biases, as supported by how local thresholding outperforms global thresholding.

Further, we qualitatively show how different thresholding schemes globally affect the connections across the whole brain. We observed that our LT has more intra-subnetwork connections and reasonable inter-subnetwork connections compared to GT at different densities, as shown in Figure 3.13.

3.2.4 Discussion

To tackle the noise reduction problem in brain network edge building, we propose a Local Thresholding scheme to deal with false positives present mostly in the FC estimates. Our technique accounts for local variations in the estimate of the connectivity strength, i.e., region bias across the connectome. Using an indirect evaluation, we attained higher accuracy in subnetwork extraction using our proposed Local Thresholding compared to conventional GT and state-of-the-art Local Thresholding (MST-KNN).

We will focus on building a direct validation method for our Local Thresholding scheme, e.g., testing on a noise model built on the ground truth of the connectivity matrices. At the same time, we need to design validations for the false positive connections discovered in tractography-based AC estimates [185].

3.3 Summary

Towards improving the brain edge building, we explored noise reduction direction due to the lack of explicit noise model. We proposed two strategies to tackle the two most common estimation errors, i.e., false negatives and false positives.

Specifically, to combat false negatives in the connectivity estimation, we proposed a matrix completion based technique by recovering missing connections. Based on the assumption that brain is comprised of a small number of subnetworks, we format the recovering missing connection problem as a low rank matrix completion problem. We effectively propose an information aggregation approach to tackle the problem that there is not a clear change point to estimate the rank due to the noisy nature of neuroimaging data. Further, we advise a neighborhood information filling strategy to solve the negative entry problem. Based on synthetic data and real HCP data, we quantitatively demonstrated the superiority of our matrix completion based approach as compared to existing methods such as median filter or simple matrix completion approaches, in terms of the recovery accuracy and the IQ prediction. The refined AC estimation derived based on our approach could benefit clinical investigations on anatomical connections on patient data with diseases, such as glioma, multiple sclerosis, and amyotrophic lateral sclerosis.

On the other hand, we present a local thresholding method to suppress false

positive connectivity estimates. Our method is able to tackle the regional bias problem, which is caused by region size bias, motion artifacts, and signal dropouts due to susceptibility artifacts. Compared to widely used global thresholding and state-of-the-art local thresholding method, our method achieved better performance in an application of subnetwork extraction in favor of a better denoising process.

Chapter 4

Graphical Metric Guided Subnetwork Extraction

This chapter focuses on improving brain subnetwork extraction using graphical metrics, which is based on the second part of papers [P5] and the first part of [P6].

The brain network can be abstracted as a graph. One can gain insights into the fundamental architectures and functions of the brain from its modular structure [192], which can be extracted by clustering brain nodes into subnetworks (also referred to communities) using community detection methods [106]. However, reliable subnetwork extraction from either the AC or FC remains challenging, mostly since the brain network topology may be obscured by noisy connectivity estimates. Even with the techniques we proposed in Chapter 3 for noise reduction, the subnetwork extraction is still not ideal. We have noticed that few subnetwork extraction methods have exploited intrinsic properties of brain networks other than modularity. In this chapter, we investigate whether incorporating a greater number of graphical metrics, which was mentioned in Section 1.4.2.4, would give us more domain-related information about the brain networks for better subnetwork extraction. At the same time, we will explore if we can build a model that would resemble the biological nature of brain subnetworks.

4.1 Modularity Reinforcement Subnetwork Extraction

4.1.1 Related Work

Functional Magnetic Resonance Imaging is widely used for studying the functional modular structure of the brain. Due to false positives and negatives in connectivity estimates and the region-specific biases, introduced in Section 3.2.1, existing subnetwork extraction methods [106] (community detection methods, graphical partitioning methods, ICA based methods) still can not generate reliable subnetwork extraction results. Different from the existing work, we aim to utilize more related graphical information into the subnetwork extraction process. Combined with the local thresholding which we proposed in Section 3.2.2, we here propose a modularity reinforcement strategy for improving brain subnetwork extraction.

4.1.2 Modularity Reinforcement Model

Our modularity reinforcement strategy is based on a similarity measure. The underlying assumption is that node pairs belonging to the same subnetwork presumably connect to a similar set of brain regions, i.e., have similar connection fingerprints. We derive a node similarity measure from the thresholded graph by comparing the adjacency structure of each node pair. We note that we apply the local thresholding proposed in Section 3.2.2 to first reduce the noise and regional bias. Given the thresholded graph presented as the adjacency matrix \mathbf{A} , where $\mathbf{A}_{i,:}$ is the connection fingerprint of node i , we define the similarity between a pair of nodes (i, j) as the number of common adjacent nodes they share, normalized by the minimum node degree of the node pair:

$$\mathbf{S}_{i,j} = \frac{\sum_{k=1}^n \mathbf{A}_{i,k} \mathbf{A}_{j,k}}{\min(d_i, d_j)}, \quad (4.1)$$

where $d_i = \sum_{k=1}^n \mathbf{A}_{i,k}$. We use the minimal degree for normalization, instead of e.g., the average degree, so that connections associated with hub nodes (nodes with more edges, defined in Section 1.4.2.4) will not be overly down-weighted. Since nodes within a subnetwork are expected to share more adjacent neighbors than nodes belonging to different subnetworks, \mathbf{S} boosts the within-subnetwork edges while

suppresses the between-subnetwork edges, which highlights the modular pattern inherent in the local thresholded connectivity matrix $\hat{\mathbf{C}}$: Hence, we use \mathbf{S} to refine $\hat{\mathbf{C}}$ to reinforce its modular structure:

$$\hat{\mathbf{C}}_{ij}^{\mathbf{S}} = \mathbf{S}_{i,j} \hat{\mathbf{C}}_{ij}. \quad (4.2)$$

4.1.3 Subnetwork Extraction Based on Graph Cuts

For subnetwork extraction, we employ Ncuts, chosen due to its wide use by the fMRI community. To set the number of subnetworks, M , we adopt an automated technique based on the spectral properties of the graph Laplacian: $\mathbf{L} = \mathbf{D} - \mathbf{W}$, where \mathbf{W} is a distance matrix derived from the connectivity matrix \mathbf{C} using Equation 2.1, $\mathbf{D}_{ii} = \sum_{j=1}^n \mathbf{W}_{i,j}$. Specifically, an eigenvalue of 1 has been shown to correspond to the transition where single isolated nodes would no longer be declared as a subnetwork [193]. We thus set M to the number of eigenvalues of \mathbf{L} with values less than 1. Setting M based on the conventional approach of modularity maximization [194] was also performed for comparison.

4.1.4 Experiments

4.1.4.1 Materials

We used the exact same synthetic and real data for local thresholding, described in Section 3.2.3.1 to validate our modularity reinforcement strategy. We additionally performed a reproducibility test using another atlas, HO atlas, which has 112 ROIs with a greater coverage of the brain than Will90fROI atlas.

4.1.4.2 Results

We compared our strategy (LTMR-LT with modularity reinforcement) against GT, LT, GT with modularity reinforcement (GTMR) and MST-KNN in [22]. LT was implemented using our proposed scheme (Section 3.2.2). GTMR was implemented by deriving adjacency matrices with global thresholding, and subsequently executing our proposed modularity reinforcement strategy (Section 4.1.2). Instead of using a specific threshold, we examined a range of graph densities to test the ro-

bustness of our proposed strategy. For synthetic data, evaluation was based on the accuracy of subnetwork extraction. To estimate accuracy, we matched the extracted subnetworks to the ground truth subnetworks using Hungarian clustering with the DSC, defined in Section 3.2.3.2. The average DSC over matched subnetworks was taken as accuracy. For real data, we assessed the overlap between the extracted subnetworks and fourteen well-established brain systems [23] and subnetwork reproducibility based on HO atlas for a range of graph densities. In the overlapping test, DSC was calculated between the estimated subnetwork membership and the given one from the well-established brain systems [23] after matching the label using Hungarian assignment [160]. In the reproducibility test, DSC was calculated between different subnetwork results on different graph densities.

4.1.4.2.1 Synthetic Data An example of the various steps of our strategy is shown in Figure 4.1 c-f to demonstrate how our strategy highlights the modular structure of the graph. With GT (Figure 4.1 c), node 2 was isolated from subnetwork 1. In contrast, our LT scheme (Figure 4.1 d) was able to preserve node 2. Also, with our LT (Figure 4.1 d), one of between-subnetwork edges (i.e., edges between nodes 6 and 7 & nodes 6 and 9) was pruned, which would help prevent the two subnetworks from being declared as one, whereas none of between-network edges was pruned using GT (Figure 4.1 c). Further, refining the graph (Figure 4.1 c-d) with our similarity measure helped to highlight the modular pattern (Figure 4.1 e-f), e.g., the between-network edges which were similar to or higher than some within-network edges (especially those edges between nodes 12 and 13, node 2 and 1 & nodes 2 and 5) in Figure 4.1 c-d were suppressed by our similarity to be the lowest values in Figure 4.1 e-f.

On the 100 synthetic dataset with 100 nodes over a density range of [0.005, 0.5] at an interval of 0.01, LTMR achieved significantly higher accuracy (average DSC = 0.6735 ± 0.0475) than GT (average DSC = 0.6216 ± 0.0417 , $p = 7.56e-10$), LT (average DSC = 0.6537 ± 0.0479 , $p = 2.89e-7$), and MST-KNN (average DSC = 0.6327 ± 0.0732 , $p = 7.38e-8$) based on Wilcoxon signed rank test with Bonferroni correction, Figure 4.2. LTMR also achieved higher DSC than GTMR (average DSC = $0.6610 \pm$, $p = 0.34$), though did not reach significance.

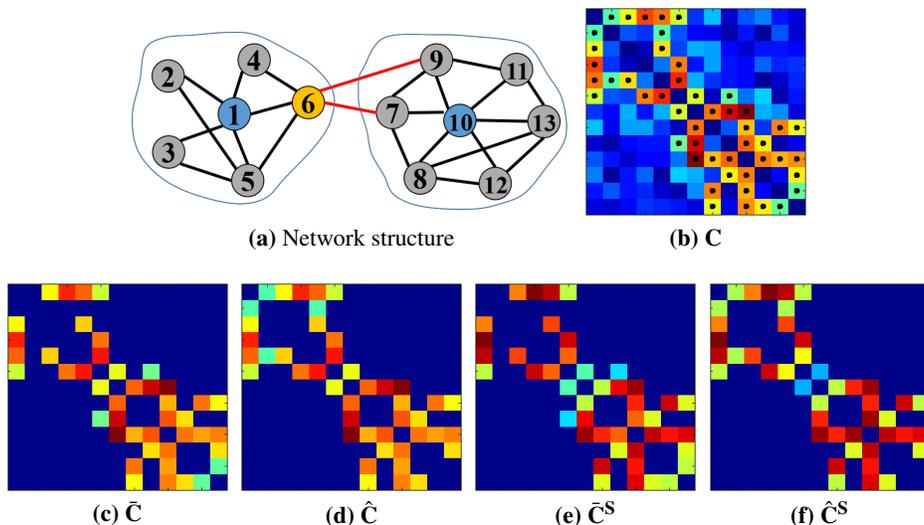


Figure 4.1: Schematic illustrating our method using small scale example having two subnetworks with each subnetwork having a provincial hub (blue) and linked by a connector hub (orange), shown in (a). In (b), warmer color indicates higher connectivity and black dots indicate the ground truth adjacency matrix. We denote \tilde{C} as global thresholded, and \hat{C} as local thresholded connectivity matrix. At a graph density of 0.25, the GT generated isolated node 2 in (c), while our LT preserved two edges linked to node 2 in (d). Refining the graph (c) and (d) suppressed the between-network edges (edges between nodes 6 and 7 & nodes 6 and 9) to be the lowest connectivity in (e) and (f).

4.1.4.2.2 Real Data We first evaluated our strategy by examining the overlap between our extracted subnetworks and 14 well-established brain systems presented in [23], which we used as ground truth, Figure 4.4 a. For this assessment, we only considered connectivity matrices based on the Will90fROI atlas [23]. Our proposed LTMR achieved an average DSC of 0.6222 ± 0.0474 , which was significantly higher than GT (average DSC= 0.5384 ± 0.0436 , $p=0.002$), MST-KNN (average DSC= 0.4567 ± 0.0653 , $p=0.002$), GTMR (average DSC= 0.5422 ± 0.0470 , $p=0.006$) based on Wilcoxon signed rank test with Bonferroni correction, and higher than LT (average DSC= 0.5936 ± 0.0374 , $p=0.063$), as shown in Figure 4.3 a. We note that an average M of 11 was estimated with the Laplace approach,

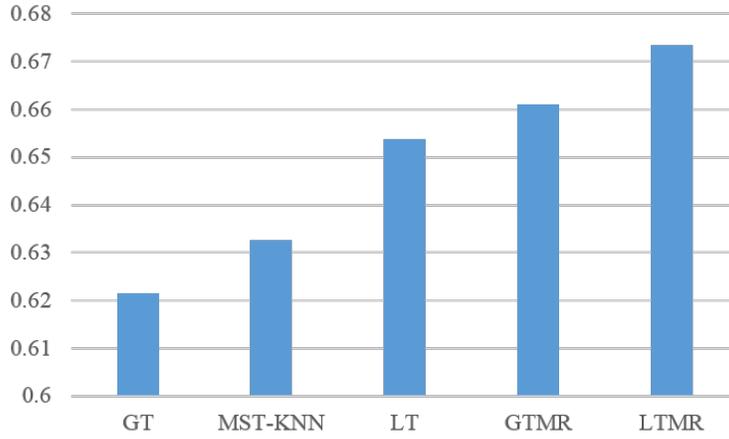


Figure 4.2: Subnetwork extraction on synthetic data at graph densities from 0.005 to 0.5 at interval of 0.01. Our proposed strategy attained the highest DSC overall.

whereas an average M of 4 was estimated with modularity maximization. This result shows the resolution limits of modularity maximization [100], i.e., it tends to underestimate the number of subnetworks in favoring network partitions with groups of modules combined into larger communities. This suggests the need to explore alternative techniques for estimating the number of subnetworks.

We next evaluated the subnetwork reproducibility over a range of graph densities. We used connectivity matrices based on the HO atlas, which has larger brain coverage than the Will90fROI atlas but does not have subnetwork labels assigned to the regions. We set subnetworks corresponding to an edge density of 0.2 as the reference. Based on the Laplace approach, the optimal number of subnetworks M was found to be 11 ± 5 over the range of graph density examined. Our proposed strategy achieved an average DSC of 0.7302 ± 0.0575 , which is significantly higher than that of GT (DSC= 0.6121 ± 0.0620 , $p=0.004$), LT (DSC= 0.6677 ± 0.0599 , $p=0.027$), MST-KNN (DSC= 0.5737 ± 0.0754 , $p=0.003$) based on Wilcoxon signed rank test with Bonferroni correction, and higher than GTMR (DSC= 0.7004 ± 0.0923 , $p=0.262$), Figure 4.3 b. The results hold with other densities used as reference.

Qualitatively, based on the subnetwork extraction resulting using the HO atlas, with GT (Figure 4.4 b), we observed two subnetworks comprising only isolated

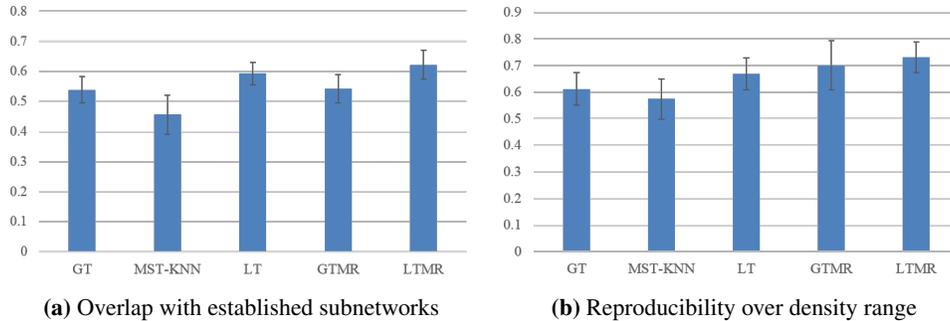


Figure 4.3: Subnetwork extraction on real data at graph densities from 0.05 to 0.5 at interval of 0.05. Our proposed strategy attained the highest DSC overall.

nodes in the left and right Pallidum (yellow and light grey node in the blue circle). We also observed that a region in the right premotor area was falsely grouped into the auditory subsystem (the light green region with a red arrow). With GTMR, two subnetworks comprising single nodes were found as well. As for LT (Figure 4.4 c), we observed the left and right insular cortex as well as the right Frontal Operculum Cortices (orange nodes with red arrows) were falsely grouped with Dorsal Default Mode regions and the left paracingulate gyrus was excluded. In contrast, our proposed strategy LTMR correctly identified known Dorsal Default Mode regions, such as paracingulate gyrus, anterior division of cingulate gyrus, and Accumbens, as a single subnetwork. Further, LT excluded the left Cuneal Cortex in the visual system (blue arrow in Figure 4.4 c). Other found subnetworks with our strategy, such as left and right executive control subnetworks (red and yellow), Figure 4.4 d, also conform well to known brain systems as was quantitatively demonstrated in Figure 4.4 a.

4.1.5 Discussion

We proposed a modularity reinforcement strategy for improving brain subnetwork extraction. By applying local thresholding in combination with modularity reinforcement based on connection fingerprint similarity, we attained higher accuracy in subnetwork extraction compared to subnetwork extraction approach with con-

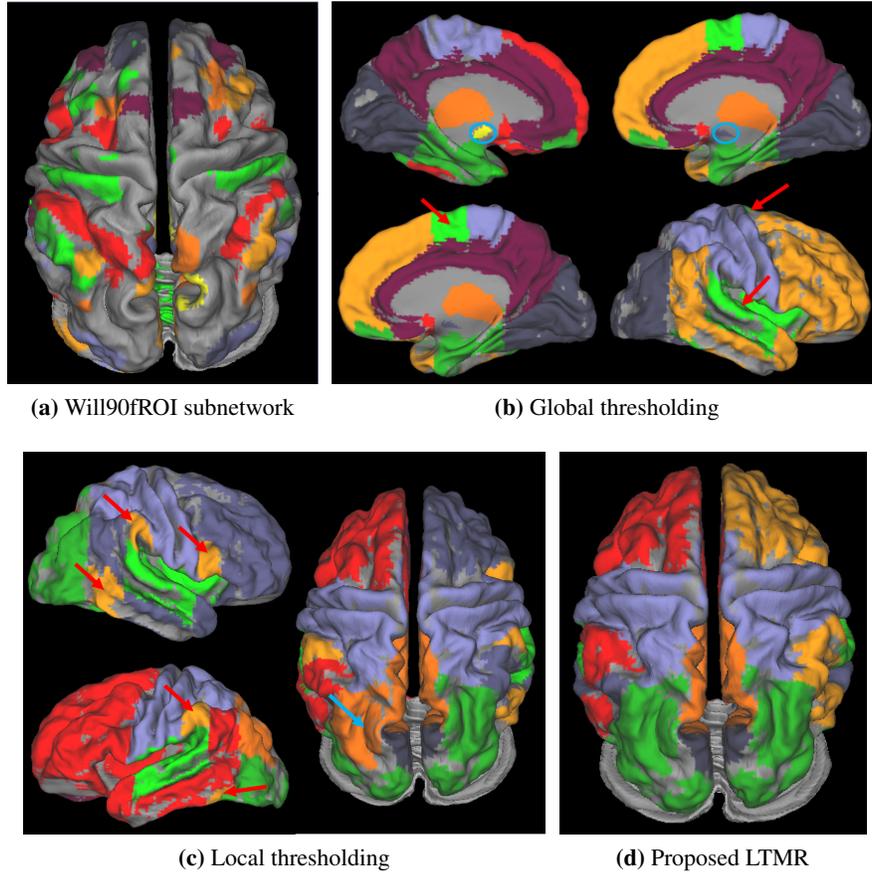


Figure 4.4: Subnetwork visualization. 11 subnetworks were extracted from graphs with a density of 0.2. (a) Well-established brain systems [23]. (b) Two subnetwork formed by isolated nodes and false inclusion of premotor-related regions into auditory system was observed using global thresholding. (c) Local thresholding failed to detect one region of known visual systems and falsely detected four unrelated regions into dorsal default mode system. (d) Our strategy LTMR correctly detect most of the subnetworks found in [23].

ventional global thresholding, the state-of-art local thresholding, and our own local thresholding. Higher overlap with established brain systems and higher subnetwork reproducibility were also shown on the real data. Our results thus demonstrated clear benefits of incorporating a greater number of graphical information with our strategy for subnetwork extraction. In fact, our strategy can be extended to applications beyond subnetwork extraction. The rescaled correlation values are not intended to be directly used for analysis, but we can derive features based on the extracted subnetworks, e.g., within-subnetwork connectivity computed from the original connectivity estimates, and use those features for group analysis and behavioural association studies. Also, we can use, e.g., mutual information or Jaccard index, to compare the cluster labels of two groups and use permutation to assess significance of group differences.

4.2 Provincial Hub Guided Random Walker Based Subnetwork Extraction

Following the section above, we are to explore if incorporating a greater number of graphical metrics would give us better subnetwork extraction. Also, we will build a model inspired by the brain subnetwork’s biological nature.

4.2.1 Related Work

Existing brain subnetwork extraction methods typically use modularity, introduced in Section 1.4.2.5 as a fitness measure to optimize a graph partitioning [106]. Modularity as defined by the Q value below reflects the intra and inter subnetwork connection structure of a network.

$$Q = \frac{1}{4m} \sum_{ij} (\mathbf{A}_{i,j} - \frac{k_i k_j}{2m}) (s_i s_j + 1) = \frac{1}{4m} \sum_{ij} (\mathbf{A}_{i,j} - \frac{k_i k_j}{2m}) s_i s_j. \quad (4.3)$$

Asides from the limitations such as resolution limit, such subnetwork extraction remains challenging due to the pronounced noise in neuroimaging data. Besides, few methods have exploited intrinsic properties of brain networks other than modularity.

Informative network metrics such as hubs, within-module degree scores, and

participant coefficients can be estimated given a graph [17]. However, these network metrics have not been used to guide the subnetwork extraction process. Previous studies on anatomical [105, 106] and functional networks [195, 196] suggest the presence of the “provincial hubs” [197], hubs that are highly connected to nodes within a subnetwork. These provincial hubs are thought to be responsible for the formation and stability of the subnetworks [197, 198]. Given the critical role of provincial hubs, we argue that incorporating provincial hubs to guide the subnetwork extraction process would be beneficial.

4.2.2 Resemblance Between Provincial Hubs and the Seeds

We here propose a Random Walker (RW) [199] based approach which utilizes brain network properties to guide the brain subnetwork extraction. We hypothesize that the manner in which the nodes are clustered into groups, based on the probabilities of walking to seeds in the RW, closely resembles the mechanism whereby brain regions within a subnetwork are inter-linked via provincial hubs. Through an iterative optimization process, we update the RW model architecture based on the feedback from the brain properties estimated from the previous iteration. The feedback mechanism enables the incorporation of network information within iterations to efficiently identify improved subnetwork assignments. RW [199] captures the probability transition along the pathway from a node to a seed, rather than just the distance between nodes and seeds in other seeded based methods such as region growing and kmeans clustering. The RW with prior model [200] is in fact well suited to incorporate a feedback from network properties via prior edges connecting nodes to augmented seeds. We thus deploy a feedback informed optimization model based on RW with prior model. Further, most clustering methods produce hard subnetwork assignments, which forces all nodes to be assigned into single certain subnetworks. By using the RW, we can infer the probability of a node being assigned to subnetworks, and further investigate the significance of nodes belonging to the subnetworks.

4.2.3 Overview of Random Walker Model

Given a weighted graph G , a set of weighted edges $w_{ij} \in E$, and a set of N nodes V , comprising labeled nodes (i.e., seeds) V_S , and unlabeled nodes, V_U , such that $V_S \cup V_U = V$, we wish to assign each node $v_i \in V_U$ with a label from set $\{1, 2, \dots, M\}$. We define the set $M_S = m^1, m^2, \dots, m^M$ as subnetwork structure, in which each subset $m^k, k \in 1 \sim M$ of M_S is a set of nodes within subnetwork k . The RW approach [199] assigns to each unlabeled node the probability, x_i^k , that a random walker starting from that node i first reaches a seed assigned to label k . Each unlabeled node is then assigned to the label for which it has the highest probability, i.e., $y_i = \max_k x_i^k$. The minimization of $x^k \top \mathbf{L} x^k$ yields the probability x^k , where $\mathbf{L}(N \times N)$ is the Laplacian matrix of the graph defined as:

$$\mathbf{L}_{v_i v_j} = \begin{cases} d^i & \text{if } i = j, \\ -w_{ij} & \text{if } v_i \text{ and } v_j \text{ are adjacent nodes,} \\ 0 & \text{otherwise.} \end{cases} \quad (4.4)$$

where $d^i = \sum_j w_{ij}$. By partitioning the Laplacian matrix into labeled $\mathbf{L}_S(M \times M)$ and unlabeled $\mathbf{L}_U(N - M \times N - M)$ blocks:

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_S & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{L}_U \end{bmatrix} \quad (4.5)$$

and denoting an $|V_S| \times 1$ indicator vector as $f_i^k = \begin{cases} 1 & \text{if } y_i = k \\ 0 & \text{if } y_i \neq k \end{cases}$, the minimization of $x^k \top \mathbf{L} x^k$ with respect to x_U^k is given by the linear equation as:

$$\mathbf{L}_U x_U^k = -\mathbf{B} f^k. \quad (4.6)$$

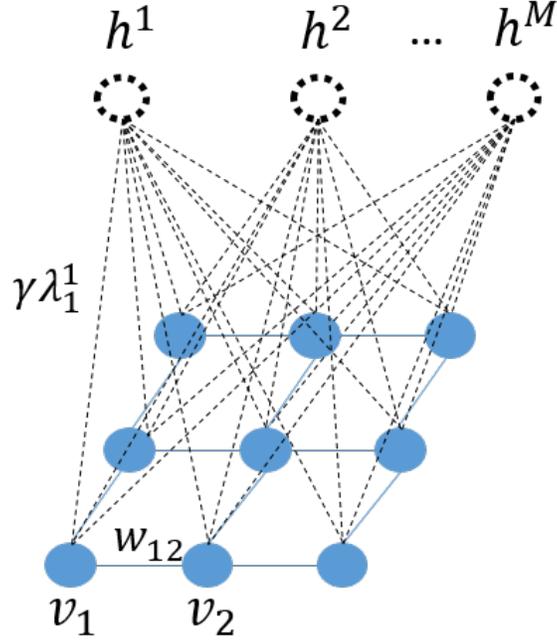


Figure 4.5: Augmented graph model for our augmented RW with prior model. The use of prior is equivalent to using M labeled “floating” augmented nodes (dash-line circles in black) that correspond to each label and are connected to each blue node with black dash lines.

4.2.4 Feedback Informed Optimization Model

4.2.4.1 Relating to RW with Prior Model

Given the important role of provincial hubs in forming and stabilizing the subnetworks [198], we propose an iterative optimization RW model by introducing an prior edge based on the RW posterior probability to reflect the affinity between a node and a seed. This affinity resembles how nodes are connected to provincial hubs. We deploy the RW with prior model [200] with an exact closed-form posterior probability update as below:

$$(\mathbf{L}_U + \gamma \sum_{r=1}^M \Lambda_U^k) x_U^k = \lambda_U^k + \mathbf{B} f^k, \quad (4.7)$$

where λ_i^k is a prior, that represents the probability of node i belonging to subnetwork k , and Λ^k is a diagonal matrix with the values of λ^k . According to [200], the incorporation of priors in Equation 4.7 yields the same solution as would be obtained for the RW probabilities on an augmented graph shown in Figure 4.5. In this augmented graph, we denote the set $H = \{h^1, h^2, \dots, h^M\}$ as the augmented seeds (black “floating” seeds indicated as dotted circles in black), which are generated by the subnetwork extraction result from the previous iteration. The blue edges in Figure 4.5 are node-to-node edges w_{ij} , and we define the black dash lines, $\lambda_i^k, k \in 1 \sim M$ as the “prior edges” connecting the nodes to the floating augmented seeds. Here γ is used to modulate the degree of the incorporation of the prior. The corresponding Laplacian matrix in our proposed model is:

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_S & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{L}_U + \gamma \mathbf{I} \end{bmatrix}, \quad (4.8)$$

where $\mathbf{B}_{i,k} = -\gamma\lambda_i^k$, \mathbf{I} is an identity matrix, and the element in the right corner unlabeled block is derived using the following equation:

$$d_{\mathbf{L}_U}^i + \sum_{k=1}^M -\mathbf{B}_{i,k} = d_{\mathbf{L}_U}^i + \gamma \sum_{k=1}^M \lambda_i^k = d_{\mathbf{L}_U}^i + \gamma, \quad (4.9)$$

where $\sum_{k=1}^M \lambda_i^k = 1$, as the probability of a node walking to all seeds sum to unity.

4.2.4.2 Multi-seed Model

Based on the augmented graph model (one floating augmented seed per subnetwork), we extend our model to contain multiple augmented seeds per subnetwork. The reason for using multiple seeds is the possible over-parcellation, which naturally splits a provincial hub into multiple seeds [201]. Further, using multiple seeds can increase robustness of the model against the problems caused by noisy connections, which could affect the power of single seed by creating strong “false” walks (e.g., through the connector hubs). By using the multi-seed strategy, we can preserve several reliable connections to battle against false positive connections (see an illustration in Figure 4.6). We empirically observed the outperformance of using multiple seeds compared to single seed.

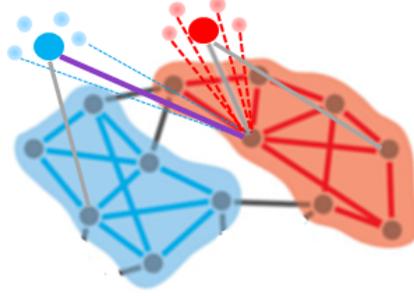


Figure 4.6: Illustration of robustness of using multiple seeds by mitigating the problems caused by noisy connections between subnetworks. Noisy connections between subnetworks could affect the power of single seed by creating strong “false” walks (thick purple line connecting the blue seed to a node which should be in the red subnetwork). By using multi-seeds (the smaller floating dots around the big dot in the center), we can preserve several reliable connections (the dash lines connecting those smaller dots to the nodes) to battle against possible false positive connections.

We denote a new set of seeds, $HM = \{H^1, H^2, \dots, H^M\}$, where each subset $H^k = \{h_1^k, h_2^k, \dots, h_{n_k}^k\}$ represents the set of multiple seeds that belong to subnetwork $k, k \in \{1, 2, \dots, M\}$. We define the prior edge connecting a node i to one of the augmented seeds h_j^k in the subnetwork k as the split version of the prior edge connecting a node to a single augmented seed in (Equation 4.10):

$$\mu_i^{h_j^k} = w_{i, h_j^k} \lambda_i^k, j \in 1 \sim n_k, \quad (4.10)$$

accordingly, the corresponding \mathbf{B} will be updated as $\mathbf{B}_{i, h_j^k} = -\gamma \mu_i^{h_j^k}$. Then we generate the subnetwork assignment probability by summing up the probabilities of a node walking to each of the multiple seeds.

4.2.4.3 Subnetwork Assignment Confidence Based on Posterior Probability

One major gain in using RW model is that it offers a confidence value that a given node belongs to a particular subnetwork (as represented by the probability). Thus the posterior probability derived from the RW model provides users with options

to discard those non-significant nodes with low probabilities belonging to any sub-networks.

4.2.4.4 Framework of the Feedback Informed Optimization Model

We first initialize the process by conducting a pre-partitioning of the brain graph. In the iterations, we update the seeds based on the subnetwork assignments from the previous iteration, and the corresponding prior edges using the probabilities of a node belonging to a certain subnetwork. We automatically set seeds based on the network properties of provincial hubs, which have high within-module degree Z score and low participation coefficient P score [17], defined as:

$$Z_i = \frac{d^i(m_i) - \bar{d}(m_i)}{\sigma^{d(m_i)}}; P_i = 1 - \sum_{m \in M_S} \left(\frac{d^i(m)}{d^i}\right)^2, \quad (4.11)$$

where m_i is the subnetwork containing node i , $d^i(m_i)$ is the within-module degree of i , i.e., the number of links between i and all the other nodes in m_i , and $\bar{d}(m_i)$ and $\sigma^{d(m_i)}$ are the corresponding mean and standard deviation of the within-module m_i degree distribution. We stress the prior edges by γ when relatively reasonable graph partitioning has been achieved from the previous iteration, based on relatively high Q value and vice versa.

The stopping criteria of the optimization on the seeds and corresponding prior edges is the convergence of the subnetwork extraction results, which is measured by the Normalized Mutual Information (NMI) [202] between the subnetwork assignments from two successive iterations.

The initial seeds are set from a pre-partitioned subnetwork structure using Ncuts [171]. We have chosen Ncuts due to its wide use in brain graph study community and its global optimality guarantees, making Ncuts outperform other partitioning methods mentioned in Section 2.1.3. We derive the prior edges in the first iteration using the average weights between a node and all the remaining nodes within a particular subnetwork, followed by a normalization to guarantee $\sum_{k=1}^M \lambda_i^k = 1$. The average weights have been chosen to increase robustness to noise.

We note that we set negative values to zero in the connectivity matrix \mathbf{C} , given

the currently unclear interpretation of negative connectivity [93]. Further, we map the connectivity matrix \mathbf{C} to the graph weights w using a Gaussian kernel, where σ is the decaying parameter which is estimated by averaging non-zero values of the connectivity distance:

$$w_{ij} = \exp\left(-\frac{(1 - \mathbf{C}_{ij})^2}{2\sigma^2}\right). \quad (4.12)$$

4.2.5 Experiments

We evaluated our proposed provincial hub guided RW based approach by comparing our approach against conventional subnetwork extraction method which does not utilize multi-pronged graphical metrics nor the biological intuition. We have chosen Ncuts due to its wide use in the subnetwork extraction studies and its merits which were mentioned in Section 2.1.3. We applied the subnetwork extraction approaches on the resting state function connectivity data.

4.2.5.1 Materials

4.2.5.1.1 Synthetic Data We generated 200 synthetic datasets that cover a wide variety of network configurations to simulate FC graphs using the technique from [203]. Each dataset comprised $N = 500$ regions and 4 scans of $t = 1200$ time points as in the real data. We set the number of subnetworks, M , to a random value between 10 and 20 in each dataset. The number of ROIs in each subnetwork was set to $\lceil N/M \rceil + q$ with q being a random number between -2 and 2, and ROIs were randomly assigned to subnetworks. We generated a $N \times N$ adjacency matrix, Σ , which was taken as the ground truth based on the network configurations. Next, we built a $\Sigma_{\mathbf{F}}$ matrix, by randomly setting $p\%$ of the values in Σ to 1 to model how FC estimates are prone to false positives. p was randomly chosen from [0, 20]. Time series were then generated by drawing random samples from $\mathcal{N}(0, \Sigma_{\mathbf{F}})$. Then Gaussian noise was added to the time series with SNR randomly set between -6 and -3 dB. Finally, FC matrices were simulated by computing the Pearson's correlation of these noisy time series.

4.2.5.1.2 Real Data We used the Resting State Functional Connectivity based on MRI scans of 77 unrelated healthy subjects from the HCP dataset [146]. Two sessions of rs-fcMRI with 30 minutes for each session were available. Preprocessing applied was described in Section 1.8.3.1. We then used the Willard atlas [204] which has 499 ROIs to define the brain nodes. We chose Willard atlas since it has subnetwork labels for 142 significant nodes belonging to 14 well-established subnetworks [23], which we used as subnetwork label pseudo ground truth. The remaining 357 nodes can be studied to verify if our RW based approach can detect the nodes with low probabilities being assigned to any subnetworks. Voxel time courses within ROIs were averaged to generate region time courses. The region time courses were demeaned, normalized by the standard deviation, and concatenated across subjects. The Pearson’s correlation values between the region time courses were taken as estimates of FC matrices.

4.2.5.1.3 Parameter Setting We examined brain graphs at different graph densities from the range [0.005, 0.5] at an interval of 0.005 to test the robustness of our approach and the local thresholding proposed in Section 3.2.2 was used. The number of seeds needed within each subnetwork is dependent on the graph density. Denser graphs require more seeds to tackle the noisy connection problem. Empirically, we set n_k to 15 for density [0.005, 0.1], 20 for density [0.105, 0.3], and 25 for density [0.305, 0.5]. For the optimization stopping criteria, we empirically set the threshold for the NMI of subnetwork assignments between successive iterations to 0.99. All the results in our experiments have reached convergence.

4.2.5.2 Results

All statistical comparisons are based on the Wilcoxon signed rank test with significance declared at an α of 0.05 with Bonferroni correction.

4.2.5.2.1 Synthetic Data We assessed our method and contrasted method by computing the DSC between the ground truth and the estimated subnetwork labels. Our proposed provincial hub guided RW based approach achieved statistically significantly higher DSC at 0.9868 ± 0.060 than the contrasted method at $0.9703 \pm$

0.046.

4.2.5.2.2 Real Data We evaluated our approach by examining the overlap between our extracted subnetworks and well-established brain systems comprising 14 subnetworks [23] which we took as the pseudo ground truth. We show that our proposed approach achieved statistically significantly higher average DSC (0.5542 ± 0.056) over a range of graph densities as compared to the contrasted method (0.5125 ± 0.054).

We further verified that the posterior probability derived from our RW based approach match with the assignment of significant and non-significant nodes in the Willard atlas. Based on an exemplar brain graph which achieved the highest subnetwork extraction DSC=0.6738 to the established brain systems [204]. Amongst the 94 matched nodes out of the 142 significant nodes in the Willard atlas, the average posterior probabilities of a node belonging to the assigned subnetwork is 0.6411, and an average value at 0.4864 probability has been found in the remaining 357 non-significant nodes.

4.2.6 Discussion

We explored utilizing intrinsic properties of brain networks to incorporate a greater number of graphical metrics for improved subnetwork extraction. We proposed a provincial hub guided random walker based optimization approach. We have chosen RW based approach since the manner in which the nodes are clustered into groups based on the probabilities of walking to seeds, closely resembles the mechanism whereby brain regions within a subnetwork are inter-linked via provincial hubs. Moreover, the posterior probability of a node belonging to subnetworks provides a way to study the significance of a node.

We have explained the outperformance of using multiple seeds over a single seed within each subnetwork. It still remains as a challenge to determine the number of seeds automatically. For simplicity, we set the number of the seeds to be the same across different subnetworks based on graph densities; however, an adjustable number should be determined based on the size of the provincial hubs and the subnetwork size. Our future work on this open question will be identifying the

fully connected subgraphs which are derived from a provincial hub. Specifically, the major provincial hub will be detected based on the brain network properties, and a clique extraction based on the major hub will generate the nodes in the fully connected subgraphs as the multiple seeds.

4.3 Summary

In order to better extract the brain subnetwork structure, we have exploited intrinsic properties of brain networks and incorporated a greater number of graphical metrics other than modularity into the process. At the same time, we built a model inspired by the brain subnetwork's biological nature.

Specifically, we proposed a modularity reinforcement strategy based on a connection fingerprint similarity concept. We attained higher accuracy in subnetwork extraction on synthetic data, higher overlap to well-established brain systems and higher reproducibility on the real HCP data compared to conventional state-of-the-art community detection methods.

We also proposed a provincial hub guided random walker based model inspired by the similarity between how the nodes are clustered into groups based on the probabilities of walking to seeds and the mechanism of brain regions within a subnetwork are inter-linked via provincial hubs. We devised a multi-seed strategy to tackle the noisy connection problem. We fully utilized informative module-related network metrics such as hubs, within-module degree scores, and participant coefficients to guide the brain subnetwork extraction. We have demonstrated that incorporating domain-related information and building a biological intuition inspired model result in better subnetwork extraction in terms of subnetwork extraction accuracy on the synthetic data and the overlaps to well-established brain systems on the real data.

Chapter 5

Multimodal/Multisource Brain Subnetwork Extraction

This chapter focuses on further improving brain subnetwork extraction based on multimodal/multisource fusion technologies, which is based on the second part of [P6], [P7] and [P8].

First, along the line of subnetwork extraction methods in the previous chapter, we further explore multimodal fusion techniques to improve subnetwork extraction. We propose multimodal provincial hub guided approach to fuse AC and FC. Moreover, we propose a high order relation informed approach based on hypergraph to combine information from rs-fcMRI and t-fcMRI for subnetwork extraction. We further study multisource approach for overlapping brain subnetwork extraction using canonical network components, i.e., cliques, which we defined as co-activated node groups across multiple tasks. Based on the overlapping subnetwork extraction results derived from co-activated cliques, we study the subnetwork overlaps and their relationships to network measures such as hubs.

5.1 Multimodal Random Walker based Subnetwork Extraction

5.1.1 Related Work

Currently, fMRI and dMRI are the most widely used modalities to estimate FC and AC for the brain subnetwork extraction purpose. We argue that combining FC and AC could help improve subnetwork extraction, as mentioned in Section 1.4.3. Most existing multimodal subnetwork extraction methods aimed to fuse the connectivity matrices from different modalities to estimate common patterns, such as representative MultiView Spectral Clustering (MVSC) [85], and related co-training methods [205–207]. MVSC achieves the agreement between two views by projecting the affinity matrix of one view to the eigenspace of the other view [85]. Other models added in techniques such as regularization: CO-training with REGularization (COREG) [208] or overlapping subnetwork assumption: Coupled Stable Overlapping Replicator Dynamics (CSORD) [203]. COREG enforces the view-specific eigenvectors to look similar by regularizing them towards a common consensus (centroid based co-regularization), and optimizes for individual clusterings as well as the consensus using a joint cost function [208]. CSORD is a multimodal integration technique based on a sex-differentiated formulation of replicator dynamics [203].

Although studies have shown the close correspondence between anatomical and functional connectivities and subnetworks [105, 113, 114], network analyses have so far failed to demonstrate a clear one-to-one correspondence between network communities in AC and FC [142]. Despite that, recent studies suggest that highly connected brain regions are greatly involved in establishing network-wide communication, acting as the focal points in large-scale anatomical and functional networks [142]. A computational model of the large-scale structure of the cerebral cortex suggested a partial correspondence between anatomical and functional hubs even at very short time scales (msecs. to secs.) [209].

Here we intend to facilitate multimodal integration, but not by trying to find the common patterns in the connectivity matrices as the existing methods, instead by using important hubs to guide the subnetwork extraction process. Meanwhile,

the limitations of existing multimodal methods include the lack of exploitation of brain properties other than modularity and the mostly hard/crisp subnetwork assignments. Based on our model, we incorporate a greater number of brain properties from the graphical metrics and provide the probabilistic membership of a node belonging to subnetworks, rather than just a crisp assignment.

5.1.2 Multimodal Provincial Hub Guided Random Walker Model

In Section 4.2, we proposed an provincial hub guided RW based model by introducing an prior edge based on the RW posterior probability, which resembles the mechanism whereby brain regions within a subnetwork are inter-linked via provincial hubs. In the model, we update the seeds based on the subnetwork assignments from the previous iteration, and the corresponding prior edges using the probabilities of a node belonging to subnetworks using Equation 5.1 and Equation 5.2.

$$(\mathbf{L}_U + \gamma \sum_{r=1}^M \Lambda_U^k) x_U^k = \lambda_U^k + \mathbf{B} f^k, \quad (5.1)$$

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_S & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{L}_U + \gamma \mathbf{I} \end{bmatrix}, \quad (5.2)$$

Inspired by the findings of close relationship between hubs in anatomical and functional networks (Section 5.1.1), we argue that the provincial hubs should be used to integrate the information flow between the two modalities. We extend the proposed feedback informed iterative optimization model in Section 4.2 to enable multimodal integration based on the provincial hub guided RW model (henceforth referred to as multi-modal Random Walker (mmRW)). The underlying assumption is that subnetworks captured via multiple imaging modalities might share information through provincial hubs. By alternating the connectivity modalities for the node-to-node edges within the iterative model in Equation 5.1, provincial hubs can be used to guide the feedback information propagation across modalities until the subnetwork assignments from different modalities converge, as measured by NMI. As illustrated in Figure 5.1, when the graph weights in the previous iteration $t - 1$ are defined by FC shown in blue edges, the weights in the current iteration t are defined by AC shown in red edges. Similarly to the unimodal RW framework (Sec-

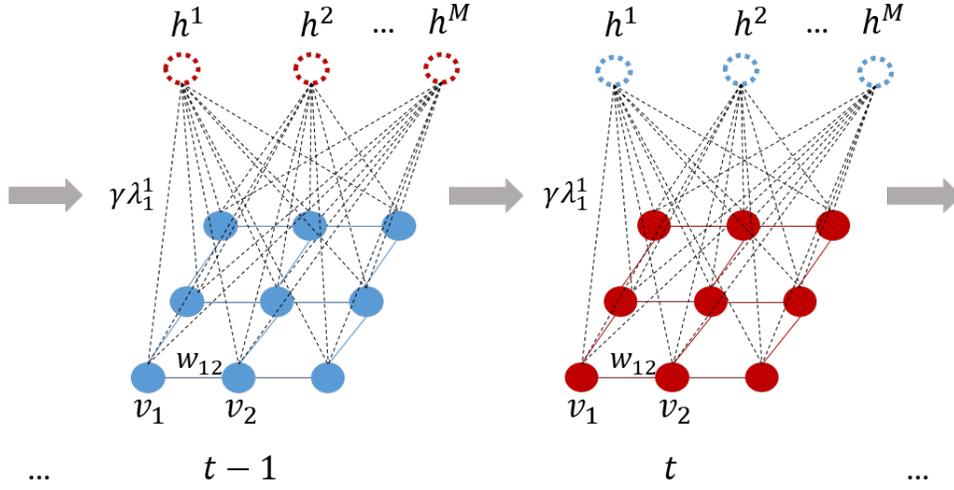


Figure 5.1: Schematic illustration of mmRW, where the graph weights in the previous iteration $t - 1$ are defined by FC edges in blue and the seeds are found using the provincial hubs derived from AC (shown as red dash circles), then the weights in the current iteration t are defined by AC shown in red edges and the seeds are found using the provincial hubs derived from FC (shown as blue dash circles).

tion 4.2.4.4), we stress the prior edges by γ when relatively reasonable graph partitioning has been achieved from the previous iteration, based on relatively higher Q value and vice versa.

We base our connectivity estimation on a fingerprint formulation [65], where we derive the connectivity matrices \mathbf{C} by estimating the cross-correlation between the fingerprint profiles of each brain node pair. The AC fingerprint profile for each brain node is defined as the fiber connection strength to the remaining brain nodes, namely, the number of tracts connecting two nodes normalized by the node sizes. The FC fingerprint profile is defined as the cross-correlation between the time courses of a particular node and all the remaining nodes in the brain. We also set negative values in \mathbf{C} to zero before mapping \mathbf{C} to the graph weights w using a Gaussian kernel as in Equation 4.12.

5.1.3 Subnetwork Assignment Confidence and Overlapping Subnetwork Exploration Based on Posterior Probability

Same as Section 4.2.4.3, multimodal RW based subnetwork extraction approach provides a confidence value that a given node belongs to a particular subnetwork (as represented by the probability). Thus the derived posterior probability can be studied to find significant nodes (detailed results in paragraph 5.1.4.2.2), or the potential overlapping subnetwork structure (detailed results in Section 5.3.3.4).

5.1.4 Experiments

5.1.4.1 Materials

5.1.4.1.1 Synthetic Data Similarly as in Section 4.2.5.1, we generated 200 synthetic datasets that cover a wide variety of network configurations to simulate AC and FC graphs using the technique from [203]. Each dataset comprised $N = 500$ regions and 4 scans of $t = 1200$ time points as in the real data. We set the number of subnetworks, M , to a random value between 10 and 20 in each dataset. The number of ROIs in each subnetwork was set to $\lceil N/M \rceil + q$ with q being a random number between -2 and 2, and ROIs were randomly assigned to subnetworks. We generated an $N \times N$ adjacency matrix, Σ , which was taken as the ground truth based on the network configurations. Next, we built an AC matrix, Σ_A , by randomly setting $p1\%$ of the values in Σ to 0 to model how AC estimates contain false negatives, and built a FC matrix, Σ_F , by randomly setting $p2\%$ of the values in Σ to 1 to model how FC estimates are prone to false positives. $p1$ and $p2$ were randomly chosen from $[0, 20]$. Two sets of time series were then generated by drawing random samples from $N(0, \Sigma_A)$ and $N(0, \Sigma_F)$. Then Gaussian noise was added to the time series with SNR randomly set between -6 and -3 dB. Finally, AC and FC matrices were simulated by computing the Pearson's correlation of these noisy time series.

5.1.4.1.2 Real Data We used the rs-fcMRI and dMRI scans of 77 unrelated healthy subjects from the HCP dataset [146]. Two sessions of rs-fcMRI with 30 minutes for each session, and one session dMRI data were available for multimodal

integration. Preprocessing applied was described in Section 1.8.3.1. We then used the same Willard atlas [204] which has 499 ROIs to define the brain nodes. Voxel time courses within ROIs were averaged to generate region time courses. The region time courses were demeaned, normalized by the standard deviation, and concatenated across subjects. The Pearson’s correlation values of time courses between one region and the remaining regions were taken as estimates of the FC fingerprint profiles. To compute the fiber counts between brain nodes, which were taken as the AC fingerprint profile estimates, we warped the Willard atlas onto the $b = 0$ volume of each subject. Subject-wise AC fingerprint profiles were concatenated for a group-level fingerprint profiles.

We examined a range of graph densities of [0.005, 0.5] at an interval of 0.005 using local thresholding in Section 3.2.2 to test the robustness of our approach. The number of seeds within each subnetwork n_k was set to 15 for density [0.005, 0.1], 20 for density [0.105, 0.3], and 25 for density [0.305, 0.5], same as the previous Section 4.2.5.1.

For the stopping criteria for convergence, we empirically set the threshold for NMI (from successive iterations) between inter-modality subnetwork assignments to 0.8. All the results in our experiments have reached convergence.

5.1.4.2 Results

5.1.4.2.1 Synthetic Data We compared our mmRW against our unimodal RW using Equation 4.7, unimodal Ncuts, multimodal methods MVSC [85], COREG [208], and CSORD [203]. For unimodal techniques, we reported their performance using both AC and FC estimates. On 200 synthetic datasets, we assessed each method by computing the DSC between the ground truth and estimated subnetwork labels as: $DSC = 2|L_{est} \cap L_{gnd}| / (L_{est} + L_{gnd})$, where L_{gnd} is the set of ROIs in the ground truth subnetwork, and L_{est} is the set of ROIs in the estimated subnetwork matched to L_{gnd} using Hungarian assignment [160]. mmRW achieved significantly higher DSC than each contrasted method at $p = 10^{-5}$ based on Wilcoxon signed rank test with Bonferroni correction (Figure 5.2). We note that the numbers of subnetworks were set to the ground truth subnetwork numbers for all methods for simplification.

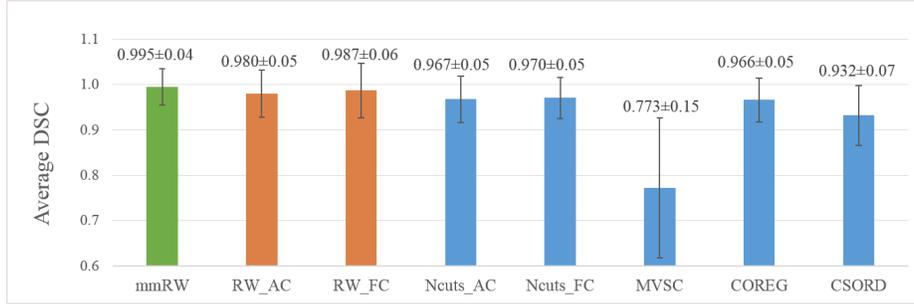


Figure 5.2: Subnetwork extraction accuracy on synthetic data. Our mmRW approach achieved significantly higher DSC than unimodal and existing multimodal methods.

5.1.4.2.2 Real Data Due to the lack of the ground truth, we first evaluated our approach by examining the overlap between our extracted subnetworks and a set of well-established brain systems [23] comprising 14 subnetworks which we took as the pseudo ground truth. We derived the connectivity matrix based on the Willard atlas [204]. Both the atlas and corresponding subnetwork assignment were manually inspected and edited by neurologists [204], which included intensive user selection. 142 out of 499 ROIs in the Willard atlas were classified as significant brain nodes, we thus compared the DSC on these significant nodes between the estimated and the established subnetwork assignments using our approach against unimodal and representative existing multimodal approaches. We set the number of subnetworks to 14 as the number of the established brain systems.

mmRW approach achieved significantly higher average DSC over a range of graph densities compared to each contrasted method at $p = 10^{-7}$ based on Wilcoxon signed rank test with Bonferroni correction, as shown in Figure 5.3. We observed that the results of using unimodal RW iterative model based on FC data can attain better results compared to AC based results. We assume that the false positives in the FC estimation can be mitigated by using our multi-seed strategy, while the false negatives in the AC estimation is harder to tackle with. The other reason might be the well-established brain systems were derived mostly by functional data. But adding AC in improved the accuracy, which confirms the benefit of multimodal integration. We note that the low DSC from CSORD could be caused by the implicit

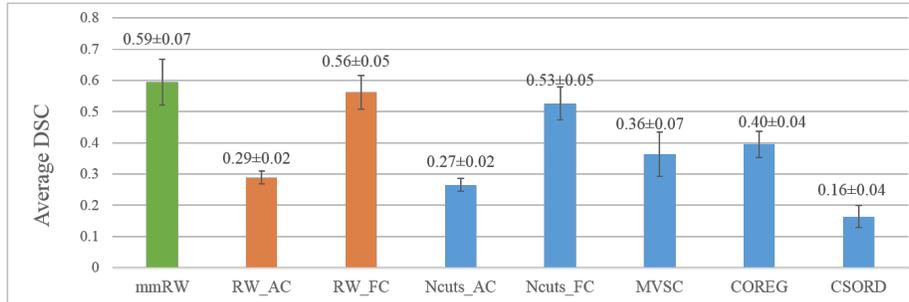


Figure 5.3: Subnetwork extraction overlap to well-established brain system [23] on real data from HCP. Our mmRW approach achieved significantly higher DSC than contrasted methods.

overlapping subnetwork assumption embedded in the approach while the pseudo ground truth has a non-overlapping subnetwork setting.

We next evaluated the inter-subject subnetwork reproducibility using the repeatability of the group subnetwork extraction results with respect to the numbers of subjects. Specifically, we compared the group subnetwork assignments generated from all 77 subjects against using only N_S subjects with N_S set between 5 and 75 at interval of 5 subjects. The average repeatability over different numbers of subjects as measured using DSC derived from our mmRW approach seems lower than our unimodal RW approach, but the difference is negligible and not statistically significant at $p = 0.5614$ for RW based on AC and $p = 0.1205$ for RW using FC based on Wilcoxon signed rank test. However, the reproducibility of our approach is significantly higher than contrasted multimodal method at $p = 10^{-2}$ based on Wilcoxon signed rank test (Figure 5.4). We note that mmRW approach only achieved comparable reproducibility compared to our unimodal RW approach since the stability was not reinforced when information was pooled from two different modalities. However, we still achieved higher reproducibility than other multimodal approaches, confirming the superior stability of mmRW amongst multimodal strategies.

In order to utilize the posterior probability derived from our approach, we selected to show the subnetwork extraction results of an exemplar brain graph which achieved the highest DSC=0.7682 to the established brain systems [204] using

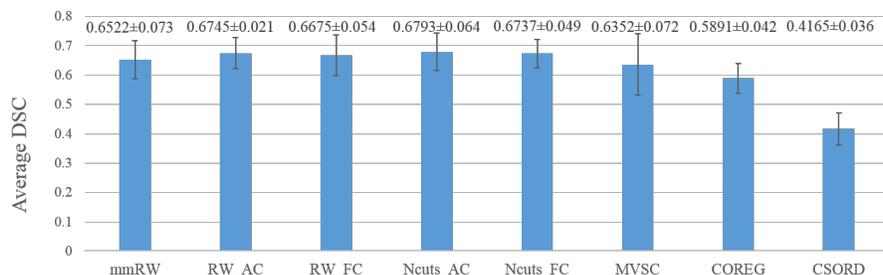


Figure 5.4: Inter-subject reproducibility on real data from HCP. Multimodal RW approach achieved comparable DSC to unimodal RW approach, but significantly higher DSC than existing multimodal methods.

mmRW. Amongst the 103 matched nodes out of the 142 significant nodes in the Willard atlas, the average posterior probabilities of a node belonging to the assigned subnetwork is 0.6473 with a minimum value at 0.3118. On the other hand, an average value at 0.4689 and a minimum value at 0.1156 of probabilities have been found in the remaining 357 non-significant nodes. This result further confirms that the nodes with higher probabilities derived from our approach matched well with those significant nodes within the 14 established brain systems. We further visualized the posterior probabilities by the size of the nodes in Figure 5.5 a-b. Here users have the option to focus on the bigger nodes by using desired thresholds of probability, or study the roles of nodes with lower probabilities, either them residing in overlapping subnetworks or simply being insignificant.

mmRW identified all commonly found subnetworks in the literature [204]. The extracted dorsal default mode subnetwork (dDMN), visuospatial subnetwork and Left Executive Control Network (LECN) are shown in Figure 5.5 c-e as exemplar results.

5.1.5 Discussion

We proposed a provincial hub guided random walker based optimization approach for brain subnetwork extraction that exploits brain properties and facilitates multimodal integration of fMRI and dMRI. Our approach helps tackle the pronounced noise in the neuroimaging data. On synthetic data, we showed that our approach achieved statistically significantly higher subnetwork identification accuracy than

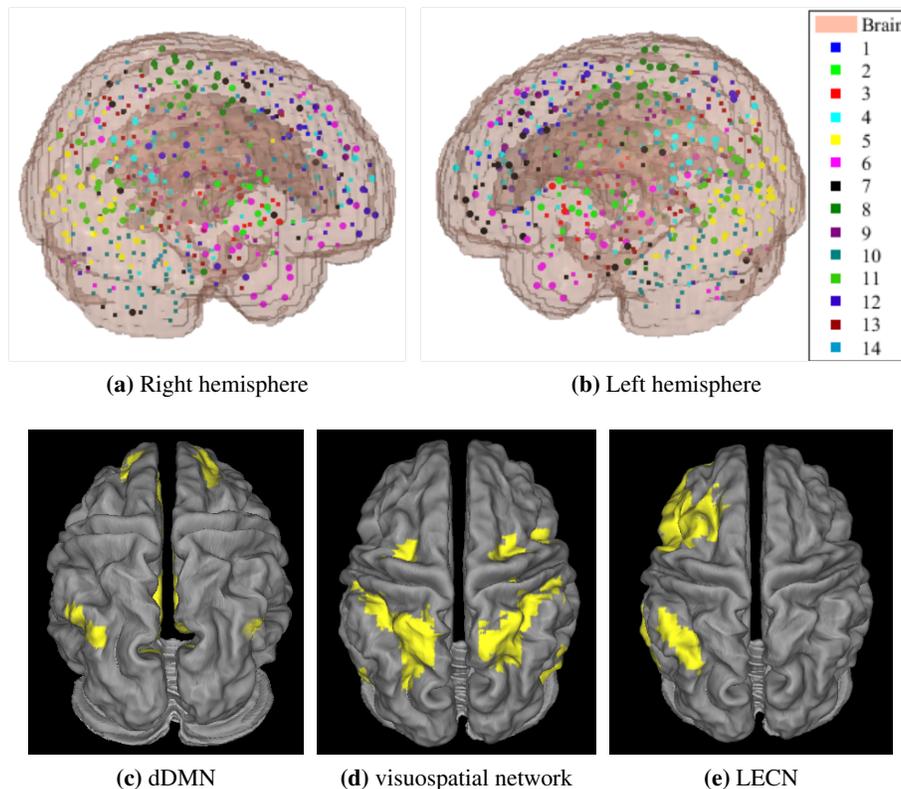


Figure 5.5: Subnetwork extraction results on an exemplar brain graph. (a-b) Probabilities of a node being assigned to a given subnetwork (color-coded in 14 subnetworks), the larger size of a node indicates a higher probability. (c-e) Exemplar subnetworks identified by mmRW.

a number of state-of-the-art approaches. On real data, we demonstrated that our estimated subnetworks matched well with established brain systems and attained comparable or higher inter-subject reproducibility. The majority of current approaches typically generate hard subnetwork assignments, while our probabilistic approach can provide users with options to examine the uncertainty of the subnetwork assignments. Our future work will focus on studying the roles of nodes with low probabilities and their possible biological meaning, e.g., them residing within overlapping subnetworks as connector hubs, or simply being insignificant to be included into any specific functional subgroups.

Our proposed multimodal integration based on RW with prior model is essentially an implicit combined loss function in terms of $Q = Q_{spatial} + \beta Q_{aspatial}$, where AC and FC are associated to one of the energy items respectively. Defining an explicit combined loss function is challenging since that there is no clear one-to-one correspondence between subnetworks from the two modalities. Instead, we used the provincial hubs to guide the integration in iterations, where the hubs have been shown, in recent studies, to reside at the focal points in both functional and structural subnetworks [142]. Our approach has been proven to outperform a directly combined loss function in the COREG method.

We used the data from 77 healthy subjects from HCP dataset Q3 release. We did not use a dataset with much bigger size since we discovered the stability at 60 subjects by examining the stability of the subnetwork assignments at the group level with respect to the numbers of subjects. Specifically, we compared the subnetwork assignments generated from all 77 subjects against using only N_S subjects with N_S being set between 5 and 75 at interval of 5 subjects. The stabilities of the subnetwork generated at the group level as measured using the DSC plateaued at 0.65 after $N_S = 60$.

One observation from our results was that the improvement of using our approach based on the synthetic data is slight, even though statistically significant. The reason could be that the noise added was not challenging enough. Gaussian noise was added to the time series with SNR randomly set between -6 to -3 dB, corresponding to the typical levels seen in task-based fMRI data, i.e., between 0.2 and 0.5. However, the approach was applied on resting state fMRI data, whose SNR is hard to determine. Our future work will be adding experiments of more noisy cases with lower SNR and more realistic percentage of missing connections based on tract tracing gold standard from the macaque data, which is similar to the humans.

Our approach does not include an automatic strategy to determine the number of the subnetworks, which is quite a challenging problem *per se*. Existing methods are largely based on finding a transition point in the matrix eigenspectrum. However, connectivity matrices typically do not display clear transition points as shown in Section 3.1. Our future work might include the multiscale version to study the hierarchy of the brain subnetworks.

5.2 Fusion of Task and Resting State Functional Connectivity for Subnetwork Extraction Based on Hypergraph

Most existing functional subnetwork extraction methods focus on resting state function connectivity data [171, 201], using functional homogeneity clustering, ICA, or graph community detection. However, resting state functional connectivity is inherently with low SNR and prone to false positive and negative correlations [90]. Such noisy resting state functional connectivity information leads to unreliable subnetwork extraction results. Given the resemblance between resting state and task functional subnetworks [111] and high order nodal relations reflected from multi-task data, we here aim to incorporate information from task data into the subnetwork extraction based on multilayer network. We explore if this integration can improve the subnetwork extraction by exploiting the mechanism of how groups of nodes collaborate together to execute a function and how these groups communicate with each other.

5.2.1 Related Work - Relationship between Task and Resting Functional Connectivity

Recent studies indicate that resting state functional activity actually persists during task performance [210], and similar network architecture is present across task and rest, which is supported by the existence of similar multi-task FC and resting-state FC matrices that were averaged across subjects [211]. Studies have also shown that there is a strong resemblance between rest and task subnetworks [106, 111]. The spatial overlap between resting-state functional subnetworks and task-evoked activities has been discovered [112, 212].

Based on the close relationship between the two, resting state data have been used to predict the task activities, by using group ICA to discover repertoires of canonical network components that will be recruited in tasks [213]; by applying the graphical connectional topology of brain regions at rest to predict functional activity of them during task [212]; or based on a voxel-matched regression method to estimate the magnitude of task-induced activity [214].

On the other hand, aggregating brain imaging data from thousands of task re-

lated studies allowed the construction of ‘co-activation networks’, whose major components and overall network topology strongly resembled functional subnetworks derived from resting-state recordings [19, 46, 215].

It has been suggested that networks involved in cognition are a subset of networks embedded in spontaneous activity [111, 216], and a number of canonical network components in the pre-existing repertoires of intrinsic subnetworks are selectively and dynamically recruited for various cognitions [213, 217].

5.2.2 Related Work - Multilayer Brain Network Analysis

Multilayer network has recently been used to model and analyze complex high order data, such as multivariate and multiscale information within the human brain [218]. Different layers can represent relationships across different temporal variations [219], reflect different imaging modalities (such as task and rest) [218], or different frequency bands [220], etc. Hypergraph is a type of multilayer graphs, in which edges can link any number of nodes [221]. Hypergraphs have been used to identify non-random structure in structural connectivity of the cortical microcircuits [222], identify high order brain connectome biomarkers for disease prediction [223], and study relationships between functional and structural connectome data [224].

5.2.3 High Order Relation Informed Subnetwork Extraction

Our assumption is that task data can be beneficial for subnetwork extraction since the repeatedly activated nodes in different tasks could be the canonical network components in the spontaneous resting state subnetworks. At the same time, the multilayer structure of repeatedly activated nodes across multi-task can be elegantly presented as a hypergraph. We propose a high order relation informed subnetwork extraction model, which (1) facilitates multisource integration of task and rest data for subnetwork extraction, (2) utilizes information from the relationship between groups of activated nodes across different tasks, and (3) enables the study on higher order relations among brain network nodes.

5.2.3.1 Framework

We propose a high order relation informed approach based on hypergraph to integrate both resting state and task information for brain subnetwork extraction. We firstly construct a brain graph based on a certain parcellation atlas. Secondly, we detect activation of brain nodes from task data to define the nodes for multiple layers in the hypergraph, and define the connection strength between nodes using task-induced connectivity. Thirdly, we construct the multitask hypergraph and incorporate resting state FC strength information when setting the weights of hyperedges. Fourthly, we fuse task and rest FC using weighted combination model before performing graphcut on the constructed graph.

5.2.3.2 Notation Overview of Hypergraph

5.2.3.2.1 Notations We here follow most of the notations presented in [221]. Let V denote a set of nodes, and E denote a family of subsets e of V such that $\cup e \in E = V$. Then we define $G = (V; E)$ a hypergraph with the vertex set V and the hyperedge set E . A hyperedge containing just two nodes is a simple graph edge. A hyperedge e is said to be incident with a node v when $v \in e$. Two nodes are *connected* if they both belong to the same hyperedge. Two hyperedges are connected if the intersection of them is not an empty set, $e_i \cap e_j \neq \emptyset$. Given an arbitrary set X , let $|X|$ denote the cardinality of X . A hypergraph G can be represented by a $|V| \times |E|$ incidence matrix \mathbf{H} with entries $h(v, e) = 1$ if $v \in e$ and 0 otherwise, see an example in Figure 5.6. A weighted hypergraph, $G = (V; E; w)$, is a hypergraph that has a positive number $w(e)$ associated with each hyperedge e , called the weight of hyperedge e . Next, we define four important measures of hypergraph properties.

For a hyperedge $e \in E$:

1. We follow [221] to define its degree as $d_h(e) = \delta(e) := |e|$, which counts the number of nodes that exist in the hyperedge. If one uses the incidence matrix, $\delta(e) := \sum_{\{v \in V\}} h(v, e)$. Let \mathbf{D}_e denote the diagonal matrices containing the hyperedge degrees. Take Figure 5.6 as an example, $\delta(e_1) = 3$, and $\delta(e_2) = 2$.

2. We further define the hyperdegree of a hyperedge as the number of hyperedges connected to it, denoted as $d_{hH}(e) := \sum_{\{e_i \in E, e_i \neq e\}} e \cap e_i$. For example,

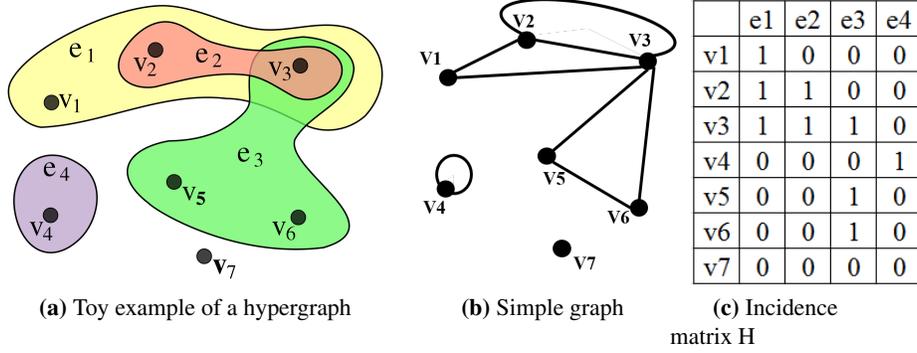


Figure 5.6: Hypergraph and its corresponding simple graph and incidence matrix. Left: an hyperedge set $E = \{e_1, e_2, e_3, e_4\}$ and a node set $V = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$. Middle: the corresponding simple graph. Right: the incidence matrix \mathbf{H} of the hypergraph on the left, with the entry (v_i, e_j) being set to 1 if v_i is in e_j , and 0 otherwise.

$d_{\mathbf{H}}(e_1) = 3$, $d_{\mathbf{H}}(e_3) = 2$, and $d_{\mathbf{H}}(e_4) = 0$ in Figure 5.6.

For a node $v \in V$:

3. We follow [221] to define its degree by $d(v) = \sum_{\{e \in E | v \in e\}} w(e)$. If one uses the incidence matrix, $d(v) = \sum_{\{e \in E\}} w(e)h(v, e)$. When all $w(e) = 1$, $d(v)$ counts the number of hyperedges which include this node: $d(v) = \sum_{\{e \in E | v \in e\}} 1$, or $d(v) = \sum_{\{e \in E\}} h(v, e)$. Let \mathbf{D}_v denote the diagonal matrices containing the node degrees.

4. We then define the hyperdegree of a node as $d_{\mathbf{H}}(v) := \sum_{\{v \in e | e \in E\}} \delta(e)$, which counts the number of nodes connected to a particular node across all hyperedges. For example, $d_{\mathbf{H}}(v_2) = 5$, $d_{\mathbf{H}}(v_3) = 6$, $d_{\mathbf{H}}(v_5) = 3$ in Figure 5.6. Its weighted version will be estimating the strength between the connected node pairs.

Next, let \mathbf{W} denote the diagonal matrix containing the weights $w(e)$ of hyperedges. Correspondingly, the adjacency matrix \mathbf{A} of hypergraph G is defined as:

$$\mathbf{A} = \mathbf{H}\mathbf{W}\mathbf{H}^T - \mathbf{D}_v, \quad (5.3)$$

where \mathbf{H}^T is the transpose of \mathbf{H} .

5.2.3.2.2 *Graphcut of the Hypergraph* One can group the nodes into subsets using graph partitioning methods, i.e., graphcut. The intuition is to find a partition of the graph such that the edges within a subset have high weights (strong intra-class connections), and the edges between different subsets have low weights (weak inter-class connections). Let $S \in V$ denote a subset of nodes and S^c denote the complement of S . Follow the notations in [157], the adjacency matrix $\mathbf{A}(X, Y) := \sum_{i \in X, j \in Y} a_{ij}$. For a given number M of subsets, the Mincut approach [225] implements the graphcut by generating a partition S_1, \dots, S_M which minimizes

$$\text{cut}(S_1, \dots, S_M) := \frac{1}{2} \sum_{i=1}^M \mathbf{A}(S_i, S_i^c). \quad (5.4)$$

To solve the problem of separating individual nodes as a subset in Mincut, RatioCut [226] and Ncuts [227] have been proposed to encode the information of the size of a subset.

$$\text{RatioCut}(S_1, \dots, S_M) := \frac{1}{2} \sum_{i=1}^M \frac{\mathbf{A}(S_i, S_i^c)}{|S_i|} = \sum_{i=1}^M \frac{\text{cut}(S_i, S_i^c)}{|S_i|}, \quad (5.5)$$

where $|S|$ measures the number of nodes in S .

$$\text{Ncut}(S_1, \dots, S_M) := \frac{1}{2} \sum_{i=1}^M \frac{\mathbf{A}(S_i, S_i^c)}{\text{vol}(S_i)} = \sum_{i=1}^M \frac{\text{cut}(S_i, S_i^c)}{\text{vol}(S_i)}, \quad (5.6)$$

where $\text{vol}(S)$ measures the volume of S by summing over the weights of all edges attached to the nodes as $\text{vol}(S) := \sum_{v \in S} ds(v)$, and node strength $ds(v)$ is the weighted version of node degree $d(v)$.

Ncuts has been widely used in image segmentation and brain study community, since it utilizes the weight information. In the following, we show that Ncuts approach can be generalized from simple graphs to hypergraphs, which has been proven in [221].

For a hypergraph $G = (V; E; w)$, a cut is a partition of V into two parts S and S^c . A hypergraph e is cut when it is incident with the nodes in S and S^c at the same time. The *hyperedge boundary* of S is defined as $\partial S := \{e \in E \mid e \cap S \neq \emptyset, e \in E \mid e \cap S^c \neq \emptyset\}$, which is a hyperedge set consisting of the hyperedges which are

cut [221]. The definition of the volume in a hypergraph $\text{vol}(S)$ is the sum of the degrees of the nodes in S , $\text{vol}(S) := \sum_{v \in S} d(v)$. Each hyperedge is essentially a fully connected subgraph, then the edges in a subgraph is called subedges, being assigned with the same weight $w(e)/\delta(e)$. When a hyperedge e is cut, there are $|e \cap S| |e \cap S^c|$ subedges are cut. Hence, the volume of ∂S is defined by

$$\text{vol}(\partial S) := \sum_{e \in \partial S} w(e) \frac{|e \cap S| |e \cap S^c|}{\delta(e)}, \quad (5.7)$$

which is the sum of weights over the subedges being cut. By this definition, we have $\text{vol}(\partial S) = \text{vol}(\partial S^c)$. Similar to the simple graphs, Normalized hypergraph cut is to keep the high intra-class connection and low inter-class connection with a partition S_1, \dots, S_M by minimizing the cut as below:

$$\underset{\emptyset \neq S_1, \dots, S_M \subset V}{\text{argmin}} \sum_{i=1}^M \frac{\text{vol}(\partial S_i)}{\text{vol}(S_i)}. \quad (5.8)$$

5.2.3.3 Task Activation Detection - Node Definition in the Hypergraph

In order to construct the multiple layers in the hypergraph, we apply the activation detection technique on the task data to define the nodes that are contained in different hyperedges. The standard way of activation detection is to use a General Linear Model (GLM) where statistics, such as t-values, reflect the degree of the similarity between the stimulus and voxel time courses. The estimated statistics produce an activation statistics map (t-map), followed by a thresholding of the map to identify the activated voxels [228]. Due to the pronounced noise in the fMRI data, activation detection at the individual level could be inaccurate [229]. In order to derive more reliable task-induced activation, we have chosen a group activation detection over the individual based approach. First, to compute the intra-subject activation patterns, a standard GLM is applied as below [228]:

$$\mathbf{Y}^i = \mathbf{X}^i \boldsymbol{\beta}^i + \mathbf{E}^i, \quad (5.9)$$

where \mathbf{Y}^i is a $t \times N$ matrix of the task-induced fMRI time courses of N brain regions from subject i , $\boldsymbol{\beta}^i$ is a $d \times N$ activation matrix to be estimated, \mathbf{E}^i is a $t \times N$ residual

matrix, and $\mathbf{X}^i = [\mathbf{X}_{\text{task}} | \mathbf{X}_{\text{confounds}}^i]$ is a $t \times d$ matrix. \mathbf{X}_{task} is the task regressors and $\mathbf{X}_{\text{confounds}}^i$ is the confound regressors. Next, we combine the activation results across subjects to assemble a group activation map, which is used to define nodes for each layer of the hypergraph. Specifically, we apply a max-t permutation test [230] on β^i aggregated from all the subjects, which implicitly accounts for multiple comparisons and control over false detections [86]. Group activation is declared at a p-value threshold of 0.05.

5.2.3.4 Strength Informed Weighted Multi-task Hypergraph

In the beginning of Section 5.2.3, we argued that multi-task information can be presented as a hypergraph, with the hyperedges being different tasks, and the nodes in each hyperedge being the brain regions activated in a certain task. In the traditional definition of hypergraph, nodes are connected to each other binarily, i.e., the edge weights between a node pair are 1 if they are connected, or 0 otherwise. We here propose a strength informed weighted hypergraph model by incorporating the strength information from the connections between nodes. We further determine the hyperedge weight $w(e)$ using the graphical measures defined in paragraph 5.2.3.2.1.

5.2.3.4.1 Pairwise Nodal Connection Strength Estimation In order to estimate the strength of the connections between two nodes, we usually use the Pearson’s correlations between time courses from pairs of brain regions. The computation for resting state connectivity matrix \mathbf{C}^{rest} has been described in Section 1.8.3.1. To produce the task-induced connectivity matrix \mathbf{C}^{task} , we use the task-induced time course information. We follow the strategy in [211] to remove all inter-block rest periods from all regions’ time courses, before computing the pairwise Pearson’s correlations across all concatenated block/event duration time courses within a task. To keep the consistency when combining information from the nodes across different layers, we keep all the \mathbf{C}^{task} having the same dimension of $N \times N$ as the \mathbf{C}^{rest} , then set the rows and columns of non-activated nodes to zero.

5.2.3.4.2 *Proposed Strength Informed Weighted Hypergraph* We present a modified hypergraph cut criteria formulation based on Equation 5.7 to incorporate pairwise nodal connection strength information from \mathbf{C} as below in Equation 5.10. The symbol $\tilde{\cdot}$ indicates the usage of strength information.

$$\tilde{\text{vol}}(\partial S) := \sum_{e \in \partial S} \tilde{w}(e) \frac{\sum_{i \in \{e \cap S\}, j \in \{e \cap S^c\}} \mathbf{C}_{ij}^e}{\delta(e)}, \quad (5.10)$$

where $\tilde{\text{vol}}(\partial S)$ is a strength informed version of $\text{vol}(\partial S)$ in Equation 5.7, \mathbf{C}^e is the connectivity matrix derived from the task corresponding to the layer e , and $\tilde{w}(e)$ is the modified weight item in the hypergraph. We propose here to incorporate strength information from the connectivity matrix and utilize the four hypergraph measures defined in paragraph 5.2.3.2.1 to determine $\tilde{w}(e)$, whose nature is the importance of the hyperedge in the hypergraph. Based on the definition of the four hypergraph measures, we exploit their corresponding biological meanings to set $\tilde{\text{vol}}(\partial S)$ and $\tilde{w}(e)$ as below:

1. The degree of a hyperedge $\delta(e)$ counts the number of brain regions that are activated in a task. To avoid the bias of the hyperedge size, $\tilde{\text{vol}}(\partial S)$ should be normalized by $\delta(e)$.

2. The hyperdegree of a hyperedge is defined as the number of hyperedges that are connected to it. Higher value indicates that more frequently activated patterns in the brain activities exist in this hyperedge. Thus, $\tilde{w}(e)$ should be proportional to $d_{\text{hH}}(e)$, i.e., $\tilde{w}(e) \propto d_{\text{hH}}(e)$.

3. The degree of a node counts the number of hyperedges that contain this node, and the biological equivalence is the number of different tasks in which one node is activated. A node with a higher degree is similar to the definition of the connector hubs residing within different subnetworks. Hence, $\tilde{w}(e)$ should be proportional to some statistics derived from $d(v)$ of the nodes in a hyperedge e . We denote the statistics computation method as *stat* here and it can be widely used statistics such as average value (mean), median value (median) and maximum value (max). Thus, $\tilde{w}(e) \propto \text{stat}(d(v))$.

4. The hyperdegree of a node reflects the number of all other nodes that are connected to it across all layers, which equals the number of connections from

other co-activated nodes to it across multiple tasks. The biological meaning of a node with a high value coincides with the definition of hubs. Hence, $\tilde{w}(e)$ should be proportional to some statistics derived from $d_H(v)$ of the nodes in a hyperedge e , i.e., $\tilde{w}(e) \propto \text{stat}(d_H(v))$. Here, in order to incorporate strength information, we apply the weighted version of $d_H(v)$, the strength of the node $d_{Hs}(v)$ as defined in Equation 5.11, i.e., $\tilde{w}(e) \propto \text{stat}(d_{Hs}(v))$.

$$d_{Hs}(v) := \sum_{\{v \in e | e \in E\}} \sum_{u \in e} \mathbf{C}_{uv}^e, \quad (5.11)$$

where \mathbf{C}^e is the task-induced connectivity matrix for the e th task.

In order to utilize strength information and hypergraph measures, we propose the $\tilde{w}(e)$ formulation as below:

$$\tilde{w}(e) := w1 \cdot d_{hH}(e) + w2 \cdot \text{stat}(d(v)) + w3 \cdot \text{stat}(d_{Hs}(v)), \quad (5.12)$$

where $w1, w2, w3$ are free parameters to control the contributions of each measure to the hyperedge.

5.2.3.5 Multisource Integration of Rest and Task fMRI

Given the close correspondence between task and rest connectivity architecture and subnetworks, we further extend the multi-task hypergraph model to integrate rs-fcMRI information. To do that, we use \mathbf{C}^{rest} for the pairwise nodal connection strength computation in Equation 5.11 as below:

$$d_{Hs}(v) := \sum_{\{v \in e | e \in E\}} \sum_{u \in e} \mathbf{C}_{uv}^{\text{rest}}, \quad (5.13)$$

Furthermore, we explicitly combine the two sources of task and rest data for subnetwork extraction. We firstly fuse the multiple layers of the multi-task hypergraph into one single layer, and secondly combine it with a resting state connectivity layer. Given that the hypergraph cut criterion (Equation 5.7) is to evaluate the aggregated sum of the cuts across all the pairwise subedges (nodal connections) in the hypergraph, we propose to aggregate the strength information between node pairs across all the layers. To do that, we transform the multiple pairwise nodal

connections across task layers (Equation 5.10) into one single nodal connection as below:

$$\bar{\mathbf{C}}_{ij}^{\text{task}} = \frac{1}{T} \sum_{k=1}^T \frac{\tilde{w}(e^k)}{\delta e^k} \mathbf{C}_{ij}^{e^k}, \quad (5.14)$$

where the subscript $k = 1, \dots, T$ is the indicator for tasks, T is the total number of tasks available, and e^k is the hyperedge in the k th layer of the hypergraph. \mathbf{C}^{e^k} is the connectivity matrix derived using the time courses in the task k using the procedure described in paragraph 5.2.3.4.1.

We next explicitly combine the two sources by a linear weighted combination between the aggregated multi-task connectivity matrix from above (Equation 5.14) and the resting state connectivity matrix in Equation 5.15 as below:

$$\mathbf{C}^{\text{t-r}} := \gamma \bar{\mathbf{C}}^{\text{task}} + (1 - \gamma) \mathbf{C}^{\text{rest}}, \quad (5.15)$$

where γ a free parameter, which can be optimized by cross-validation, or determined by the number of the tasks available. Our linear model for combining two sources, which are both derived from functional modality, was motivated by the study indicating a largely linear superposition of task-evoked signal and resting state modulations in the brain [210]. We also explore combining the two by applying a multislice community detection approach [231], which extends modularity quality function based on the stability of communities under Laplacian dynamics with a coupling parameter ω to control over interslice correspondence of communities.

5.2.4 Experiments

We first investigated the similarity of connectivity between resting state and task-general and task-specific connectivity. To evaluate our proposed approaches, we assessed the graphical metric modularity Q value, the inter-subject reproducibility and examined the biological meaning of subnetwork assignments. We applied subnetwork extraction on (1) resting state FC alone, (2) task-induced FC alone, (3) multi-task hypergraph, (4) multi-task hypergraph integrated with resting state connectivity strength, (5) weighted combination of (4) and resting state FC, (6) combination of (4) and resting state FC using multislice community detection method

[231].

5.2.4.1 Materials

We used the resting state fMRI and task fMRI scans of 77 unrelated healthy subjects from the HCP dataset [146]. Two sessions of resting state fMRI with 30 minutes for each session, and 7 sessions of task fMRI data were available for multisource integration. The seven tasks are working memory (total time: 10:02), gambling (6:24), motor (7:08), language (7:54), social cognition (6:54), relational processing (5:52) and emotion processing (4:32). Preprocessing applied was described in Section 1.8.3.1. We then used the HO atlas [57], which has 112 ROIs, to define the brain region nodes. We chose the well-established HO atlas because it sampled from every major brain system, and consists of the highest number of subjects with both manual and automatic labelling technique compared to other commonly used anatomical atlases. We didn't use the parcellation we generated in Chapter 2 to avoid potential circularity or biases by using an independently identified node community partition. We also didn't use Willard atlas as the previous section Section 5.1, since HO has bigger coverage of the brain with comprehensive annotation of each region while Willard atlas doesn't include the annotation of the anatomical areas of each region and only 142 out of 499 nodes are assigned within the functional systems. Voxel time courses within ROIs were averaged to generate region time courses. The region time courses were demeaned, normalized by the standard deviation. Group level time courses were generated by concatenating the time courses across subjects. The Pearson's correlation values between the region time courses were taken as estimates of FC matrices. Negative elements in all connectivity matrices were set to zero due to the currently unclear interpretation of negative connectivity [93]. For task activation, we applied the activation detection on the seven tasks available following the steps described in Section 5.2.3.3.

We summarize here the annotation of the graphs for six methods being evaluated for subnetwork extraction. (1) Resting state FC matrix \mathbf{C}^{rest} is used. (2) The task general FC \mathbf{C}^{task} was generated by concatenating the time courses across all tasks before the Pearson's correlation. In (3), we use task-specific FC in Equation 5.11 and Equation 5.12 for each hyperedge, denoted as $\mathbf{C}^{\text{hyper-task}}$. We imple-

ment (4) by using resting state FC in Equation 5.11 and Equation 5.12 as described in Section 5.2.3.5, denoted as $\mathbf{C}^{\text{hyper-t-r}}$. For (5), we first generate $\bar{\mathbf{C}}_{ij}^{\text{task}}$ by using task-specific FC as $\mathbf{C}_{ij}^{e^k}$, and resting state \mathbf{C}^{rest} to compute $\tilde{w}(e^k)$ based on Equation 5.11 and Equation 5.12. We next applied our proposed local thresholding [5] on resting state FC \mathbf{C}^{rest} to match with the graph density of $\bar{\mathbf{C}}^{\text{task}}$ at 0.2765, which lies within the normal range of thresholding before subnetwork extraction between [0.2, 0.3] [171]. We then estimate $\mathbf{C}^{\text{t-r}}$ using Equation 5.15. We empirically set free parameters w_1, w_2, w_3 to one, and the *stat* to median value. For (6), we generated the $\bar{\mathbf{C}}_{ij}^{\text{task}}$ and thresholded \mathbf{C}^{rest} as the same way as in (5), then the multisource integration is implemented using a multislice approach [231], denoted as $\mathbf{C}^{\text{t-r-multislice}}$. We set the weighting for multisource integration γ or coupling parameter ω from 0.01 to 1 at an interval of 0.01. In order to perform fair comparison, \mathbf{C}^{rest} in method (1) and \mathbf{C}^{task} in method (2) have also been local thresholded at the graph density of 0.2765. Method (1) to (5) used Ncuts and (6) used generalized Louvain as the graph partitioning approach. The number of subnetworks was set to seven given that there are seven tasks available to examine if subnetwork assignments can be related to tasks. We note that setting the number of subnetworks is non-trivial as discussed in the previous section that we leave as future work. All statistical comparisons are based on the Wilcoxon signed rank test with significance declared at an α of 0.05 with Bonferroni correction.

5.2.4.2 Similarity of FC Between Resting State and Task Data

We observed a similarity at $\text{DSC} = 0.7845$ between resting state FC and task general FC, which was generated by concatenating the time courses across all different tasks. For seven specific tasks, the corresponding DSC between task-specific FC and task general FC are 0.8971 for emotion processing, 0.8557 gambling, 0.8676 for language, 0.9043 for motor, 0.8594 for relational processing, 0.8307 for social cognition, and 0.8751 for working memory. This high similarities confirms the findings in [211] that a set of small but consistent changes common across tasks suggests the existence of a task-general network architecture distinguishing task states from rest.

When resting state FC is compared to task-specific FC, the DSC are 0.7193

for emotion processing, 0.7689 for gambling, 0.7390 for language, 0.7067 for motor, 0.7533 for relational processing, 0.7659 for social cognition and 0.7118 for working memory, respectively. The variation of similarities between task-specific and resting state FC around a relatively high average level further confirms that the brain’s functional network architecture during task is configured primarily by an intrinsic network architecture which can be present during rest, and secondarily by changes in evoked task-general (common across tasks) and task-specific network [211].

These findings confirms the close relationship between task and rest, and the support for integrating multitask information into resting state based subnetwork extraction.

5.2.4.3 Modularity Q Value

Modularity Q value, defined in Equation 4.3, has been used to assess a graph partitioning through reflecting the intra- and inter- subnetwork connection structure of a network [106]. We observe that Q values of group level subnetwork extraction for method (1)-(6) are 0.1401, 0.1282, 0.1624, 0.1711, 0.2290 and 0.1905 when γ and ω were selected at the highest inter-subject reproducibility.

At the subject-wise level, the modularity Q values estimated from the subnetwork extraction using method (1)-(6) are 0.1397 ± 0.0142 , 0.1234 ± 0.0159 , 0.2072 ± 0.0199 , 0.2094 ± 0.0189 , 0.2183 ± 0.0192 , and 0.2089 ± 0.0165 respectively, Figure 5.7.

We show that the modularity estimated from subnetworks extracted based on simply concatenating task time courses is lower than using resting state data. Using hypergraph framework (3) $\mathbf{C}^{\text{hyper-task}}$ and (4) $\mathbf{C}^{\text{hyper-t-r}}$ achieves statistically higher modularity values than using either resting state data or simple concatenation of task data. Moreover, incorporating resting state information into the hypergraph framework (5) $\mathbf{C}^{\text{t-r}}$ can increase modularity compared to hypergraph method. Multislice integration (6) $\mathbf{C}^{\text{t-r-slice}}$ results in a lower modularity than (5) the linear model; however, it still outperforms all the other uni-source methods. Overall, incorporating resting state information explicitly using a weighted combination strategy, i.e., method (5) gives a statistically higher modularity than all contrasted

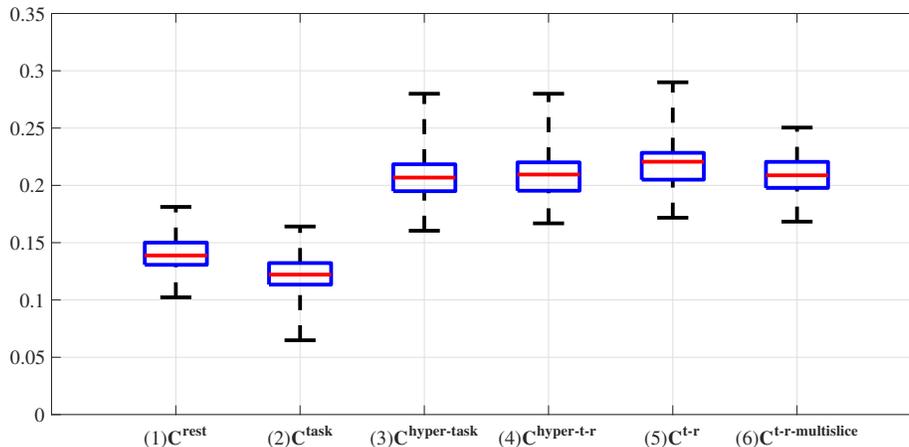


Figure 5.7: Subject-wise level modularity Q values using Method (1)-(6). For method (5) and (6), parameter γ and ω were selected at the highest inter-subject reproducibility.

methods at $p < 10^{-4}$ based on Wilcoxon signed rank test. We note that the Q values derived here are around 0.2, when the number of the subnetworks was set to seven, i.e., the number of tasks. It is relatively low due to the inherent resolution limit of Q , i.e., Q decreases when the number of subnetworks increases. We explored this direction by achieving the similar level of Q values around 0.3-0.4 when the number of subnetworks decreases to 4 as in [19].

5.2.4.4 Inter-subject Reproducibility of Subnetwork Extraction

We assessed the inter-subject reproducibility by comparing the subnetwork extraction results using subject-wise data against the group level data. The average DSC between subject-wise and group level subnetworks across 77 subjects based on methods (1)-(6) are 0.6362 ± 0.0828 , 0.5704 ± 0.0872 , 0.7083 ± 0.1094 , 0.7258 ± 0.1201 , 0.7561 ± 0.1199 , and 0.7406 ± 0.0725 , Figure 5.8. We noticed that the reproducibility using resting state FC C^{rest} is higher than simple concatenation of task time courses data C^{task} . It could be that there exist great differences in reaction to stimuli from different subjects, and simple concatenation is hard to discover the higher order relationship between canonical network components. On the other hand, analyzing multi-task information using hypergraph (3) $C^{\text{hyper-task}}$

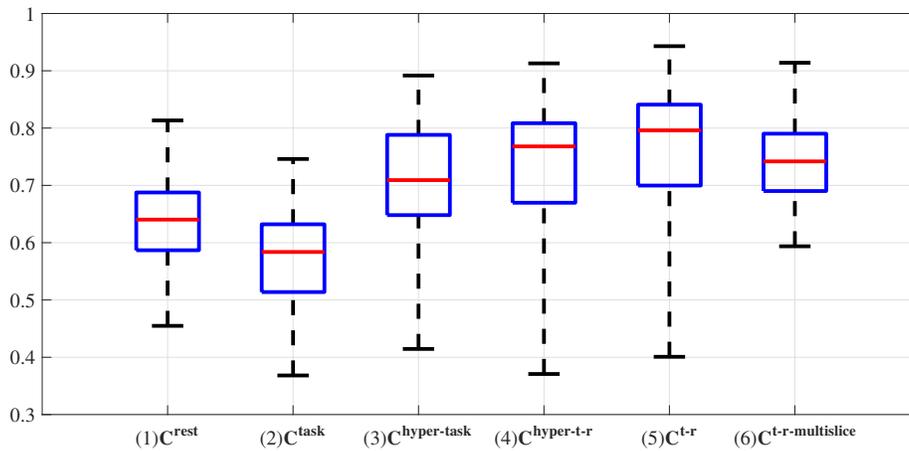


Figure 5.8: Subject-wise level inter-subject reproducibility of subnetwork extraction using Method (1)-(6). For method (5) and (6), parameter γ and ω were selected at the highest inter-subject reproducibility.

achieved much higher stability in subnetwork extraction, and incorporating resting information implicitly within the hypergraph (4) $C^{\text{hyper-t-r}}$, or explicit weighted combination (5) $C^{\text{t-r}}$ can even further enhance reproducibility. We note that the weighted combination outperforms multislice integration (6) $C^{\text{t-r-multislice}}$, which is still better than all the other uni-source methods. The reason could be that a simple linear model suffices the fusion of task and rest data. Overall, the inter-subject reproducibility derived by (5) $C^{\text{t-r}}$ is statistically higher than all contrasted methods at $p < 10^{-4}$ based on Wilcoxon signed rank test.

5.2.4.5 Biological Meaning

We next examined the biological meaning of the subnetworks extracted from method (1) - (6), where γ was set to 0.5 to report the results when resting state and hypergraph based multitask information are equally combined as an example. Seven subnetworks were extracted based on the number of tasks available. Method (1) detects most of the traditional resting state subnetworks with several false positive and negative detection. The results of method (2) oftentimes combined some important regions from different subnetworks, which lacks biological justifications. Method (3) and (4) generate similar results and both improve the results of method

(2) greatly when bringing task dynamics into the subnetwork extraction. Overall, method (5) detects brain regions, which are more biologically meaningful, by combining the intrinsic network architecture from resting state data and the task dynamics based on high-order hypergraph. We report our findings in details as the following and the visualization of subnetwork extraction results can be found in Figure 5.9.

Using method (1) based on resting state FC alone, subnetwork 1 and 6 are detected as left and right side of a combination of ECN and frontoparietal network, which include superior frontal gyrus, middle frontal gyrus, inferior frontal gyrus, posterior supramarginal gyrus, angular gyrus, frontal orbital cortex, and frontal operculum cortex. Method (1) mistakenly classified left inferior lateral occipital cortex and left anterior supramarginal gyrus into the LECN. Anterior supramarginal gyrus is part of the somatosensory association cortex, which interprets tactile sensory data and is involved in perception of space and limbs location or language processing, thus it should be included in DMN instead of ECN [232]. On the other hand, our proposed method (5) detects both the left and right sides of most of the anterior portion of ECN and posterior supramarginal gyri for subnetwork 1. Using method (5), the left inferior lateral occipital cortex was not include in ECN, which is more accurate. Besides, method (5) clustered anterior supramarginal gyrus symmetrically into subnetwork 6, which includes both sides of Posterior Cingulate Cortex (PCC), precuneus, and angular gyrus, comprising most of the posterior portion of DMN defined in [232]. As for method (2), the simple concatenation of multitask time courses, subnetwork 1 consists of frontal medial cortex and only the left side of frontal orbital cortex, and subnetwork 6 consists of most of the anterior portion of ECN, angular gyrus and only the left posterior supramarginal gyrus, which should be symmetrically included in DMN. Besides, there are two other ROIs, left subcallosal cortex and left caudate, included in subnetwork 6, which lacks biological meaning. Subnetwork 1 derived from method (3) and (4) both consist of most of the anterior portion of ECN, except that method (3) has two more one-sided frontal areas, which makes (4) more biological meaningful (with symmetric results). Subnetwork 6 of method (3) and (4) both consist of one isolate area: left anterior parahippocampal gyrus, which further indicates that there is need to incorporate resting state information into the multitask based on hyper-

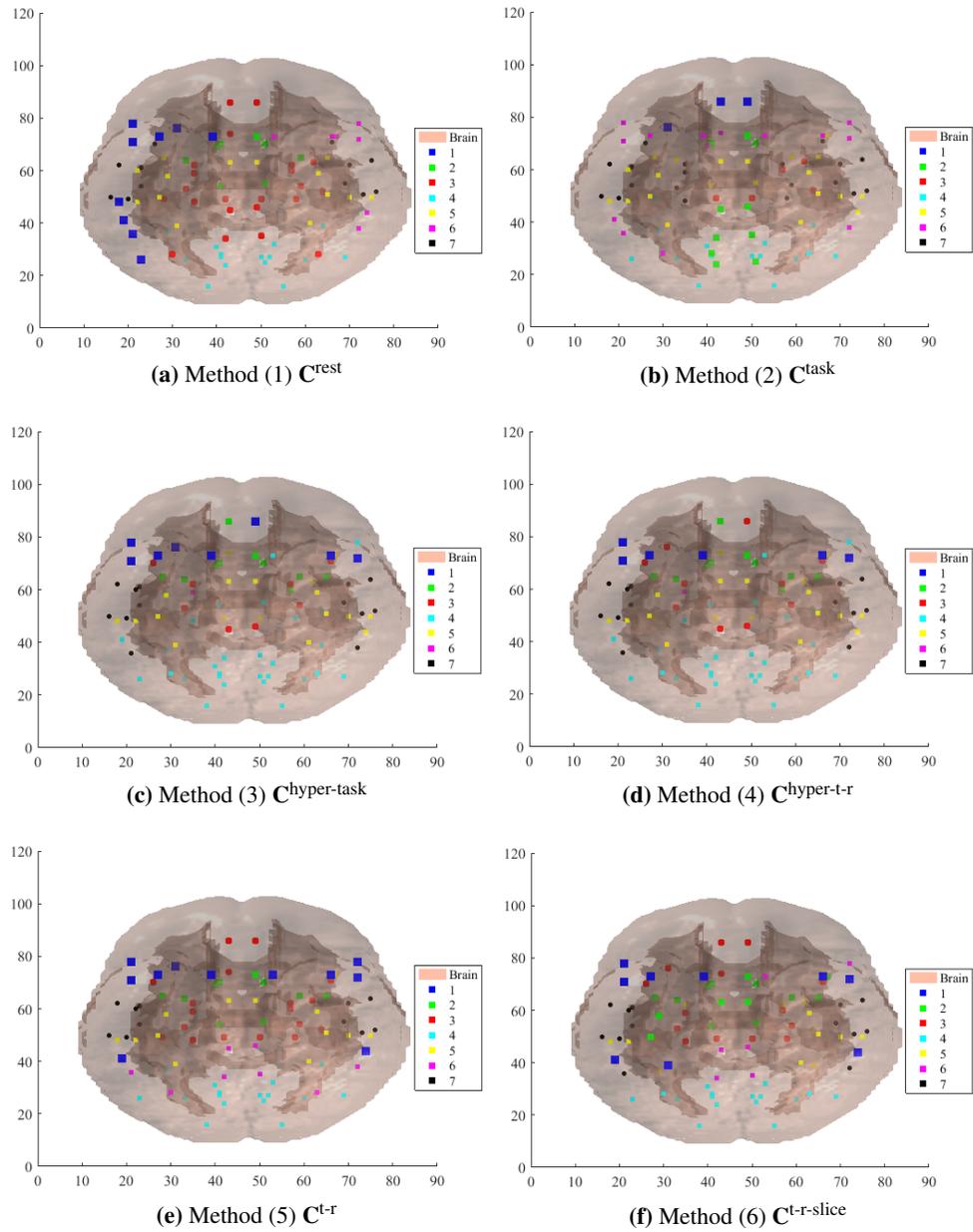


Figure 5.9: Visualization of subnetworks extraction using methods (1)-(6). The brain is visualized in the axial view. The mass center of each ROI is plotted in the MNI space and colorcoded by the membership of seven subnetworks.

graph framework.

Subnetwork 2 of method (1) includes both sides of Anterior Cingulate Cortex (ACC), caudate, thalamus, putamen and accumbens. Method (5) includes all the same brain regions as method (1) plus one other region, the insula. This subnetwork should be related to the gambling task and emotional processing, which expect to activate ACC [233, 234], ventral striatum (such as thalamus [234] and accumbens [235]), and insula [236]. Usually insula is part of the salience network and has been found to play key roles in emotional processing [237]. However, using method (1), the insula was clustered into subnetwork 5 (mostly motor system). Method (2) included right ACC and both sides of PCC, precuneous, left side of supracalcarine cortex, and accumbens inside subnetwork 2, which seems like a mixture of part of DMN, one-sided region from motor system, and one region from gambling system. As for method (3) and (4), they both extracted similar regions for subnetwork 2 as using method (5), except that they missed thalamus and falsely included left frontal medial cortex.

Subnetwork 3 derived from method (1) includes superior lateral occipital cortex, frontal medial cortex, left subcallosal cortex, PCC, precuneous, parahippocampal gyrus, temporal fusiform cortex, brain stem, hippocampus and amygdala. This assignment does not make too much sense by clustering regions from visual, auditory, emotion circuit and frontal system together. Meanwhile, the results using method (5) consists mostly of emotion circuit and social processing, which includes brain stem [238], hippocampus and parahippocampal gyrus [239], amygdala [240], and subcallosal cortex [241]. Method (5) also detected regions related to auditory functions such as temporal pole, which is reasonable since the negative emotion was induced by listening to stories. Subnetwork 3 detected by method (2) includes right anterior parahippocampal gyrus, temporal fusiform cortex and brain stem, which still lacks important brain regions in the emotion circuit. Method (3) detects more biologically meaningful regions than (2), such as hippocampus and amygdala. Using method (4) can even detect more related regions than method (3), such as frontal orbital cortex [242].

Method (1) and (5) detected almost the same brain regions for subnetwork 4, which is the visual system, except that method (5) detected one more region of left inferior lateral occipital cortex, making the results more symmetric. This

subnetwork includes inferior lateral occipital cortex, intracalcarine cortex, cuneal cortex, lingual gyrus, occipital fusiform gyrus, temporal occipital fusiform cortex, occipital pole, and supracalcarine cortex. Method (2) detected most of the visual regions except for cuneal cortex and the right supracalcarine cortex. Method (3) and (4) detected extra regions in right ECN and auditory system besides all the regions found using (5) in the visual system.

Subnetwork 5 derived from method (1) comprises of the motor system, including precentral gyrus, postcentral gyrus, only the right side of anterior supramarginal gyrus, juxtapositional lobule cortex; and the frontoparietal network including left central opercular cortex, superior parietal lobule, and parietal operculum cortex. Method (5) generated similar results as method (1), only that the results are more symmetric, which include both sides of anterior supramarginal gyrus (part of somatosensory association cortex); and more accurate in terms of frontoparietal network, which includes frontal operculum cortex instead of central opercular cortex. Both method (3) and (4) generated similar regions for subnetwork 5 as well, which includes motor system and frontoparietal network, except that they both included brain stem into this subnetwork. However, method (2) mis-classified insula, putamen and thalamus into the motor and frontal parietal networks. We note that the motor system and frontoparietal network are clustered together, it could be that the working memory tasks recruited both the motor system and frontoparietal network.

As for the subnetwork 7, both method (1) and (5) detected brain regions corresponding to language task and related auditory regions, such as anterior superior temporal gyrus, planum temporale, planum polare, and Heschls gyrus (includes H1 and H2) [243]. Different from method (1), method (5) included central opercular cortex, which can be explained by how fronto-opercular is related to language [244]. Method (2) detected some false positive brain regions in the language system such as parahippocampal gyrus, hippocampus and amygdala. Method (3) and (4) correctly clustered all the brain regions into the language network as method (5).

Method (6) generated similar results compared to method (5), only a couple regions in subnetworks 2 and 5 were switched, a couple regions in subnetwork 6 and 7 were switched, and a couple regions in 1 and 6 were switched. Overall, The subnetwork results derived by method (5) C^{t-r} have more biological meaning than

contrasted methods.

5.2.5 Discussion

5.2.5.1 Hypergraph Encodes Higher Order Nodal Relationship

Subnetwork results derived from methods based on hypergraph achieved higher modularity, higher inter-subject reproducibility, and more reasonable biological meaning than traditional connectivity analysis of pairwise correlation between nodes. These results indicate that hypergraph, which is a natural presentation of multi-task activation, can be explored to study higher order relations among the network nodes. The proposed strength informed version of automatic weight setting of the hyperedge incorporates connectivity information to reveal more accurate higher order relationship among nodes rather than just using binary information.

5.2.5.2 Multisource Integration Improves Subnetwork Extraction

We have proved that multisource integration of task and rest information can improve subnetwork extraction compared to using a single source in terms of graphical metrics, inter-subject reproducibility, along with biologically meaningful subnetwork assignments. We note that the implicit integration of rest information into multitask hypergraph achieved less improvements as the explicit integration based on the linear combination. The reason could be that the limited number of tasks available restricts the comprehensive representation of the brain using the hypergraph. Thus, by integrating rest data to compensate possible missing information resulted in overall better outcomes. Another observation is that the linear combination outperforms the multislice community detection, which still performs better than uni-source approaches. Our assumption is that rest and task FC are both derived from a single functional modality, which complements each other by revealing the two sides of FC, i.e., the resting intrinsic side and the activated evoked side. Thus, a simple linear weighted combination would suffice this situation, which outperforms other alternative combination approach in practice.

5.2.5.3 Limitations and Future Directions

There are several limitations in our present work. First, our study investigated only seven available tasks with high quality data and decent amount of data per task. This sample of seven tasks is not enough. A possible solution is to have access to both task and rest data from previous task studies or co-activation studies, which cover much wider variety of tasks. At the same time, with much more information from a greater amount of task data, we can devise a reliable automatic manner to determine the integration weighting parameter γ . The underlying rationale is that with more tasks available, we can rely more on the hypergraph based multitask source, hence the higher γ .

Secondly, we set the number of the subnetworks to be seven, which corresponds to the number of tasks available. The reason is simply to see if we can associate the subnetwork results to different tasks and gain insights from the findings based on task-induced functions. We note that setting the number of subnetworks is non-trivial as discussed in the previous chapters. Our future work will focus on exploring a finer scale of subnetwork extraction using multi-scale hierarchical approach, such as recently proposed multislice approach [211], which would improve the interpretation of the findings.

Moreover, we have explored multimodal integration of resting state FC and AC in the previous section, and we are to investigate if combining FC along with task information with AC information will further improve the multimodal subnetwork extraction. We have conducted preliminary studies on applying C^{t-r} and AC as two modalities using our proposed RW based approach. However, more systematic validation approaches and related neuroscience study evidences are required for comparing three sources with two sources.

5.3 Clique Based Multisource Overlapping Brain Subnetwork Extraction

5.3.1 Related Work - Overlapping Brain Subnetwork Extraction

The mainstream of brain subnetwork extraction and standard definition of modularity focus on nonoverlapping definition. However, studies have shown evidences of

the existence of overlapping brain subnetworks, hence the methods for nonoverlapping subnetwork extraction are limited by neglecting inclusive relationships [99]. There are emerging approaches for discovering overlapping modular network structure, which implies that single nodes may belong in more than one specific module. We here summarize some representative approaches used in brain subnetwork extraction application, and detailed information can be found in a review paper on general overlapping community detection [245].

The Clique Percolation Method (CPM) is one of the earliest methods for overlapping community detection [246]. It is based on the assumption that communities tend to be comprised of overlapping sets of cliques, i.e., fully connected subgraphs. It identifies overlapping communities by searching connected cliques. First, all cliques of a fixed size k must be detected, and a clique adjacency matrix is constructed by taking each clique as a vertex in a new graph. Two cliques are considered connected if they share $k-1$ nodes. Communities are detected corresponding to the connected components of the clique adjacency matrix. Since a vertex can be in multiple cliques simultaneously, mapping the communities from the clique level back to the node level may result in nodes being assigned to multiple communities [106, 245]. The limitation of CPM is that it operates on binarized graph edges, thus cannot handle weighted graphs [86].

A new definition of modularity has been proposed to discover the overlapping subnetwork based on unbiased cluster coefficients using resting state connectivity [99]. However, methods based on the modularity function Q suffer from degenerate partitions and resolution limit [106].

Another line of studies is to transform a network into its corresponding line graph, where the nodes represent the connections in the original network. Thus, the nonoverlapping community detection (modularity maximization used in [247] and agglomerative hierarchical clustering used in [248]) on the line graph will result in overlapping subnetworks in the original network. There exist inherent limitations in the nonoverlapping community detection used for the line graph (resolution limit for modularity maximization and local sub-optimum for hierarchical clustering).

Fuzzy community detection algorithms quantify the strength of association between all pairs of communities and nodes [245]. Fuzzy k-means clustering [249] and fuzzy affinity propagation [250] have been applied to detect overlapping brain

subnetwork extraction. However, one has to use an *ad hoc* threshold for extracting interacting nodes or independent nodes from the membership vector.

Local expansion and optimization algorithms grow a *natural* community [251] or a partial community based on local benefit functions [245]. One example is Connected Iterative Scan (CIS), which has been explored for brain subnetwork extraction [252]. Taking each node as a partial subnetwork, CIS expands the subnetwork by determining if any other nodes belong to this existing subnetwork using a local function to form a densely connected group of nodes. Its limitation is the sensitivity to a density factor that controls subnetwork size [253]. Another good example is the Order Statistics Local Optimization Method (OSLOM) [254], which uses statistical significance of a subnetwork when tested against a global randomly generated null model during community expansion. OSLOM has been shown to outperform many state-of-the-art community detection techniques.

In a previous work from our lab, the Replicator Dynamics (RD) concept from theoretical biology for modeling the evolution of interacting and self-replicating entities was used to identify subnetworks. Further, the RD formulation was extended to enable overlaps between subnetworks by incorporating a graph augmentation strategy [255]. This approach, Stable Overlapping Replicator Dynamics (SORD) [253], has demonstrated its superiority over many commonly used overlapping subnetwork extraction methods, including OSLOM.

Most of the algorithms aforementioned are based on one single source, such as resting state functional connectivity. CSORD, the multimodal version of SORD, is one of the few overlapping methods which considers multi-source information. CSORD is based on survival probabilities of different genders in evolution and graph augmentation [255]. However, its theoretical background for overlapping assumption based on graph augmentation has relatively indirect neuroscientific justifications. We here explore the direction of integrating multisource information for the overlapping subnetwork extraction by using the straightforward *clique* concept.

5.3.2 Co-activated Clique Based Multisource Overlapping Subnetwork Extraction

Recent study has indicated that repeatedly activated nodes in different tasks could be canonical network components in the pre-existing repertoires of intrinsic sub-

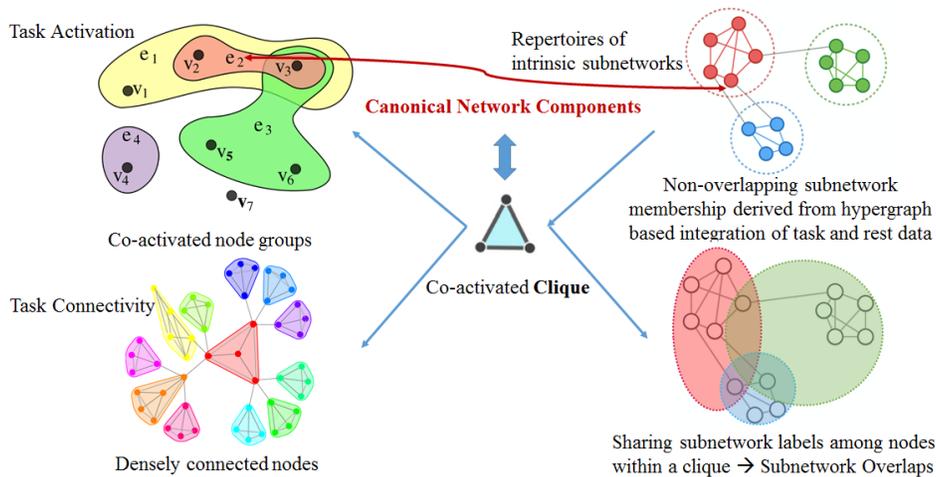


Figure 5.10: The schematic illustration of multisource clique based overlapping subnetwork extraction approach.

networks [213], we argue that the *clique* concept closely resembles groups of nodes which are the *canonical network components*. Based on the basic observation that typical communities consist of several cliques that tend to share many of their nodes [246], clique-based approach would be a straightforward way to find overlapping brain subnetworks. However, the existing clique-based subnetwork extraction approach CPM (kclique) [246] has three major limitations that it can only handle binary graphs, but not weighted graphs; the size k of cliques is fixed, which needs to be adjusted for different types of networks; and it only uses uni-source information. In order to tackle the aforementioned limitations, we here propose a multisource subnetwork extraction approach based on co-activated *clique*, which (1) uses task co-activation and task connectivity strength information for clique identification, (2) automatically detects cliques with different sizes having more neuroscientific justifications, and (3) shares the subnetwork membership, derived from multisource hypergraph based approach we proposed in the previous section, among nodes within a clique for overlapping subnetwork extraction. The schematic illustration of our approach is shown in Figure 5.10.

We first detect co-activated groups of brain nodes across different tasks based on an activation fingerprint idea, and then identify densely connected cliques based

on task-induced weighted connectivity. Core cliques are further detected using clique properties we defined. The nodes within a clique should belong to the same subnetworks due to the close relationship between nodes in a fully connected clique, we thus share the subnetwork membership of nodes within a clique to facilitate overlapping subnetwork assignment. The initial subnetwork membership for each node is derived from non-overlapping subnetwork extraction technique, which is based on the fusion of resting state connectivity and task information embedded with high order relations using hypergraph.

The difference of our approach from the traditional uni-source kclique method is the utilization of both the task co-activation information and the connectivity weights (only the binarized connectivity is used in kclique method). The co-activated cliques derived using our approach have flexible clique sizes, which has more neuroscientific justifications than the fixed k . Besides, we explore if our proposed clique node subnetwork membership sharing idea can generate more straightforward and biologically meaningful results than the existing multisource method CSORD.

5.3.2.1 Clique Identification Based on Task Co-activation

We define cliques as co-activated groups of brain nodes that are densely connected in our approach. We first identify the co-activated groups of brain nodes (coarse cliques) using an activation fingerprint idea. Then we refine coarse cliques into cliques, within which nodes are densely connected to each other based on task-induced connectivity information. We denote the clique set as CS , and the coarse clique set as CCS . Given T different tasks, one can construct a hypergraph with an $N \times T$ incidence matrix \mathbf{H} , where N is the number of brain regions and T is the number of tasks (hyperedge e). $h(v, e) = 1$ when the brain region node v is activated in the task corresponding to hyperedge e . The task-induced connectivity matrix \mathbf{C}^{task} is generated by removing all inter-block rest periods from all regions' time courses and computing pairwise Pearson's correlations of time courses which were concatenated through block/event durations across all the tasks. The underlying assumption for our clique identification is that nodes in the same clique should be co-activated across tasks at times from $t = 1 \dots T$, where t indicates the number

of tasks, in which the nodes are co-activated. There are two steps involved in our clique identification, which (1) pre-selects sets of coarse cliques in all T layers, (2) and refines the coarse cliques into cliques.

The approach starts with a pre-selection of coarse cliques CCS , which might include loosely connected nodes that are co-activated. Take each row from the incidence matrix \mathbf{H} as a activation fingerprint vector f corresponding to the task activation pattern of a node. For example, if one node is activated in the 1st, 3rd and 6th out of the seven tasks, the corresponding $f = [1010010]$. We next operate *bit-wise and* between the fingerprints from a node pair $\{i, j\}$, which gives us an output fingerprint vector of co-activation patterns \mathbf{SF}_{ij} :

$$\mathbf{SF}_{ij} = f_i \wedge f_j, \quad (5.16)$$

where f_i and f_j are the activation fingerprint vectors of node pair i and j , and \mathbf{SF} is the matrix containing the co-activation fingerprint vectors between the nodes in each node pair. We then define a matrix \mathbf{NT} which counts the number of co-activated tasks between two nodes:

$$\mathbf{NT}_{ij} = \sum_{t=1}^T \mathbf{SF}_{ij}(t), \quad (5.17)$$

where \mathbf{SF}_{ij} is the co-activation fingerprint vector of length T . Next, we define the node set $PS^{=t}$ which contains nodes that are co-activated together for t times as:

$$PS^{=t} = \bigcup_{\forall i, j \text{ s.t. } \mathbf{NT}_{ij}=t} \{i, j\}, \quad (5.18)$$

and define the node set $PS^{>t}$ which contains nodes that are co-activated together for greater than t times as:

$$PS^{>t} = \bigcup_{\forall i, j \text{ s.t. } \mathbf{NT}_{ij}>t} \{i, j\}. \quad (5.19)$$

Based on the definition above, we follow the four steps as below to identify the coarse cliques.

Step 1 We extract M_t pre-selected sets of co-activated coarse cliques from the

nodes in $PS^{=t}$. We identify $\{CCS_1^{=t}, CCS_2^{=t}, \dots, CCS_{M_t}^{=t}\}$ by ensuring all the node pairs within a certain set share the same co-activation fingerprint vector in \mathbf{SF} :

$$CCS_m^{=t} = \{p_1^m, p_2^m, \dots, p_{N_m}^m \mid \exists p_i^m, p_j^m \in PS^{=t}, s.t. \mathbf{SF}_{p_i^m p_j^m} = \mathbf{SF}_{p_1^m p_2^m}\}. \quad (5.20)$$

where $m = 1, \dots, M_t$. The minimal rank of $CCS_m^{=t}$ is 2, being only one node pair within a coarse clique. The nodes identified in a coarse clique are fully connected to each other defined by sharing the same co-activation pattern.

Step 2 Similarly, we extract M_t extended sets of co-activation coarse cliques, $\{CCS_1^{>t}, CCS_2^{>t}, \dots, CCS_{M_t}^{>t}\}$, from the nodes in $PS^{>t}$, based on the co-activation patterns between nodes in $CCS_m^{=t}$ and $PS^{>t}$:

$$CCS_m^{>t} = \bigcup_{\forall i \in CCS_m^{=t}, \exists j \in PS^{>t}, s.t. \mathbf{SF}_{ij} \wedge \mathbf{SF}_{p_1^m p_2^m} = \mathbf{SF}_{p_1^m p_2^m}} \{j\}. \quad (5.21)$$

We do not consider the coarse clique set selection for the nodes which only exist in the node set $PS^{>t}$ for the t th layer, since those will be selected in the pre-selected sets in $t + 1$ layer.

Step 3 We then generate M_t coarse clique sets by merging the pre-selected and extended sets together as:

$$\begin{aligned} CCS^t &= \{CCS_1^t, CCS_2^t, \dots, CCS_{M_t}^t\} \\ CCS_m^t &= CCS_m^{=t} \cup CCS_m^{>t}, m = 1, \dots, M_t \end{aligned} \quad (5.22)$$

Step 4 Extract the coarse clique set CCS across layers in the order from T to 1:

$$CCS = \bigcup_{t=T, \dots, 1} CCS^t. \quad (5.23)$$

The second part of clique identification is to refine the coarse cliques into cliques. When we extract $CCS = \{CCS^1, CCS^2, \dots, CCS^M\}$, there still exist loosely connected nodes in the coarse cliques, mostly from lower layers when t is small, especially when $t = 1$. Hence, we subsequently extract cliques based on the strength information from task-induced connectivity matrix \mathbf{C}^{task} and hypergraph properties. We formulate a coarse clique set, $CCS^k = \{p_1^k, p_2^k, \dots, p_{M_k}^k\}$ where there are

M_k nodes within, as an $M_k \times M_k$ simple graph with the weights between nodes being the task-induced connectivity pairwise edge strength. We next apply a local thresholding [5] on the $M_k \times M_k$ connectivity matrix $\mathbf{C}^{\text{task-}k}$ to find out the most closely connected nodes to each node, and binarize the thresholded matrix to generate an adjacency matrix $\mathbf{A}^{\text{task-}k}$. We then transform $\mathbf{A}^{\text{task-}k}$ into its hypergraph $\mathbf{H}^{\text{task-}k}$ using Equation 5.3: $\mathbf{A} = \mathbf{H}\mathbf{W}\mathbf{H}^T - \mathbf{D}_v$, where the locations with 1 in each hyperedge correspond to the nodes that comprise a fully connected subgraph, i.e., cliques CS^c . We extract N_c cliques:

$$CS = \{CS^1, CS^2, \dots, CS^{N_c}\}. \quad (5.24)$$

5.3.2.2 Clique Property Computation

We present three properties that can be derived to study the cliques for further network analysis.

(1) Co-activation times NT^c of a clique CS^c , i.e., the number of ones in the clique co-activation fingerprint:

$$CSF^c = \bigwedge_{\forall i \in CS^c} f_i, \quad (5.25)$$

then the co-activation times:

$$NCOA^c = \sum_{t=1}^T CSF^c(t). \quad (5.26)$$

(2) Activation times in a clique:

$$NA^c = \frac{1}{|CS^c|} \sum_{\forall p \in CS^c} \sum_{t=1, \dots, T} f_p(t). \quad (5.27)$$

(3) Clique overlap ratio - the times of a clique overlaps with other cliques divided by the size of a clique, i.e., the number of nodes within a clique. We first define the set of cliques which node i belongs to as a label set:

$$LC_i = \{c_1^i, c_2^i, \dots, c_{N_i}^i\}, c_k^i \in 1, \dots, N_c, \quad (5.28)$$

where LC_i is an empty set when node i does not belong to any cliques. We then define the clique overlap ratio as:

$$RCO^c = \frac{1}{|CS^c|} \left| \bigcup_{\forall p \in CS^c} LC_p \right|. \quad (5.29)$$

5.3.2.3 Core Clique Identification

Based on the clique properties, we further identify core cliques out of clique sets for the future overlapping subnetwork extraction. We argue that core cliques should have relatively high co-activation times, high activation times, and high clique overlap ratio. We then devise a core clique selection criterion based on the combination of the clique properties. We normalize all the property values into the range of $[0, 1]$ by dividing individual values by the maximum across all the cliques. The criterion is set as below:

$$\rho = \frac{\text{median}_{\forall i \in CS} \{NCOA^i\}}{\max_{\forall i \in CS} \{NCOA^i\}} + \frac{\text{median}_{\forall i \in CS} \{NA^i\}}{\max_{\forall i \in CS} \{NA^i\}} + \frac{\text{median}_{\forall i \in CS} \{ROC^i\}}{\max_{\forall i \in CS} \{ROC^i\}}. \quad (5.30)$$

For any clique c which satisfies the criterion:

$$\frac{NCOA^c}{\max_{\forall i \in CS} \{NCOA^i\}} + \frac{NA^c}{\max_{\forall i \in CS} \{NA^i\}} + \frac{ROC^c}{\max_{\forall i \in CS} \{ROC^i\}} > \rho, \quad (5.31)$$

it is selected into the core clique set.

5.3.2.4 Clique Based Overlapping Subnetwork Extraction

Based on the identified core cliques, we further deploy a subnetwork membership sharing technique to identify overlapping subnetworks. The underlying rationale is that the nodes residing within the same clique behave very similarly to perform some basic functions in tasks, thus, they should be within the same subnetworks.

In a brain graph with N nodes, let \mathbf{C}^{rest} be an $N \times N$ resting state connectivity matrix, and we have already labeled the non-overlapping subnetwork membership for each node using \mathbf{C}^{rest} . We have also defined the clique membership of a node i as LC_i in Equation 5.28. We then share the subnetwork membership of the nodes

within a clique to facilitate overlapping subnetwork assignment.

First, M_s subnetworks are extracted using non-overlapping community detection approach applied on \mathbf{C}^{rest} . We define the subnetwork membership of a node i as:

$$label(i) = s, i \in 1, \dots, N, s \in 1, \dots, M_s. \quad (5.32)$$

Next, we deploy a sharing scheme of the subnetwork membership label from $label(i)$ of a node i , with the label set of the remaining nodes in the clique where node i belongs to:

$$LS(i) = \bigcup_{\forall c \in LC_i} \bigcup_{\forall p \in CS^c} label(p), \quad (5.33)$$

and

$$label(i) = label(i) \cup LS(i). \quad (5.34)$$

We have also explored replacing the resting state connectivity matrix \mathbf{C}^{rest} with the multisource connectivity matrix $\mathbf{C}^{\text{t-r}}$ defined in Equation 5.15. We argue that we should further integrate the activation information from task data with high order relation information presented by hypergraph and the rest data when identifying the non-overlapping subnetwork membership.

5.3.3 Experiments

We first compare our multisource clique based approach against the uni-source kclique method [246], which is the closest straightforward way to identify overlapping subnetworks. Next we compare against SORD, which has been proven to outperform the state-of-the-art techniques such as OSLOM, and CSORD (the multisource version of SORD) [203] to see if our proposed approach have more direct biological intuition for the overlapping subnetwork extraction. We also examine the nodes within subnetwork overlaps derived by our approach by assessing the probability of a node belonging to subnetworks using our recently proposed multimodal RW approach [6], to verify that our overlapping subnetwork assignments correspond with the posterior probability.

5.3.3.1 Materials

Similarly as the previous section, we used the resting state fMRI and task fMRI scans of 77 unrelated healthy subjects from the HCP dataset [146]. The same pre-processings as in the previous section has been applied on resting state and task data. The resting state, task connectivity matrices, and the task activation were generated using the same procedures described in the previous section. Same as before, we applied our proposed local thresholding [5] on resting state FC \mathbf{C}^{rest} to match with the graph density of $\bar{\mathbf{C}}^{\text{task}}$ at 0.2765, which lies within the normal range between [0.2, 0.3] for thresholding before subnetwork extraction [171]. The non-overlapping subnetworks were derived from the resting state connectivity matrix \mathbf{C}^{rest} or multisource connectivity matrix $\mathbf{C}^{\text{t-r}}$ using Ncuts, when the number of subnetworks was set to 7, same as in the previous section.

We further applied local thresholding [5] on $\bar{\mathbf{C}}^{\text{task}}$ by setting graph density to be 0.1 to generate the hypergraph when we identified cliques from the coarse clique set. We selected a relatively strict threshold to only select those most closely connected nodes to form cliques. 0.1 has been chosen based on the cross-validation on inter-subject reproducibility within the range between 0.03 to 0.2 at the interval of 0.01.

We compared the overlapping subnetwork extraction using our proposed Multisource Clique-based Subnetwork Extraction (MCSE) approach with $\mathbf{C}^{\text{t-r}}$, or \mathbf{C}^{rest} against the uni-source kclique approach [246], SORD [253], which has been demonstrated to outperform state-of-the-art overlapping community detection methods including OSLOM when applied to brain subnetwork extraction, and CSORD [203], the multisource extension of SORD. Two uni-source approaches extract overlapping subnetworks using resting state data. The parameters for kclique were set using the cross-validation on the clique size k based on inter-subject reproducibility from the suggested range [3, ..., 6] [246] and reasonable graph densities from 0.03 to 0.2 at the interval of 0.01. SORD and CSORD applied 100 bootstraps by sampling with replacement as suggested in [253]. We also evaluated the probability of a node being assigned to a subnetwork using our recently proposed RW based approach [6] to examine the proposed clique-based overlapping subnetwork identification. All statistical comparisons are based on the Wilcoxon signed rank

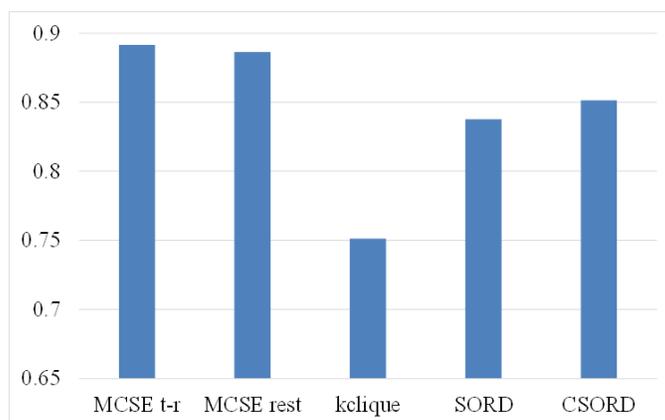


Figure 5.11: Group-level Subnetwork Extraction reproducibility based on data from two different sessions. MCSE outperforms all other contrasted approaches.

test with significance declared at an α of 0.05 with Bonferroni correction.

5.3.3.2 Comparison with Existing Overlapping Subnetwork Extraction Methods

We quantitatively evaluated the contrasted approaches based on test-retest reliability and inter-subject reproducibility, since ground truth subnetworks are unknown for the real data of human brain.

5.3.3.2.1 Group-level Subnetwork Extraction Reproducibility We first assessed the test-retest reliability based on group level subnetworks extracted separately from two sessions of rest and task data (each of the seven tasks includes two sessions of fMRI data) using DSC. The subnetworks extracted from the first session’s data are taken as the “ground truth”, against which the subnetworks from the second session are compared. We found that our proposed MCSE outperforms all other contrasted approaches, by achieving a DSC between subnetworks extracted from two sessions of data at 0.8917 with C^{t-r} and 0.8865 with C^{rest} , against kclique at 0.7514, SORD at 0.8378, and CSORD at 0.8514, see Figure 5.11.

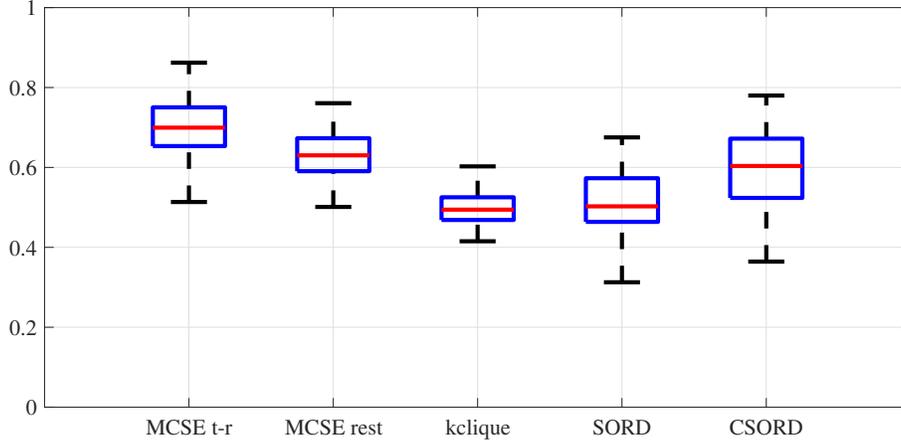


Figure 5.12: Subject-wise level inter-subject reproducibility of subnetwork extraction. Our proposed MCSE approach outperforms existing state-of-the-art overlapping community detection methods.

5.3.3.2.2 Subject-wise Subnetwork Extraction Reproducibility We assessed the inter-subject reproducibility by comparing the subnetwork extraction results using subject-wise data against the group level data, Figure 5.12. The average DSC between subject-wise and group level subnetworks across 77 subjects based on five approaches are MCSE with C^{t-r} at 0.7024 ± 0.0722 , MCSE with C^{rest} at 0.6281 ± 0.0583 , kclique at 0.4967 ± 0.0430 , SORD at 0.5129 ± 0.0774 , and CSORD at 0.5952 ± 0.0901 , respectively. MCSE with both C^{t-r} and C^{rest} are found to achieve statistically higher inter-subject reproducibility than constrained approaches based on the Wilcoxon signed rank test at $p < 10^{-10}$ and $p < 0.005$, respectively. Further, MCSE with C^{t-r} outperforms C^{rest} at $p < 0.00001$, which confirms the benefit of incorporating the task information embedded with higher order relations in assigning non-overlapping subnetwork membership.

5.3.3.3 Biological Meaning - Analyzing Function Integration

We further examined the biological meaning of the overlapping subnetworks found using all five methods, i.e., our proposed MCSE with C^{t-r} , MCSE with C^{rest} , kclique, SORD and CSORD, Figure 5.13. We first measured the overlapping ratio by dividing the number of nodes residing in the subnetwork overlaps by the total

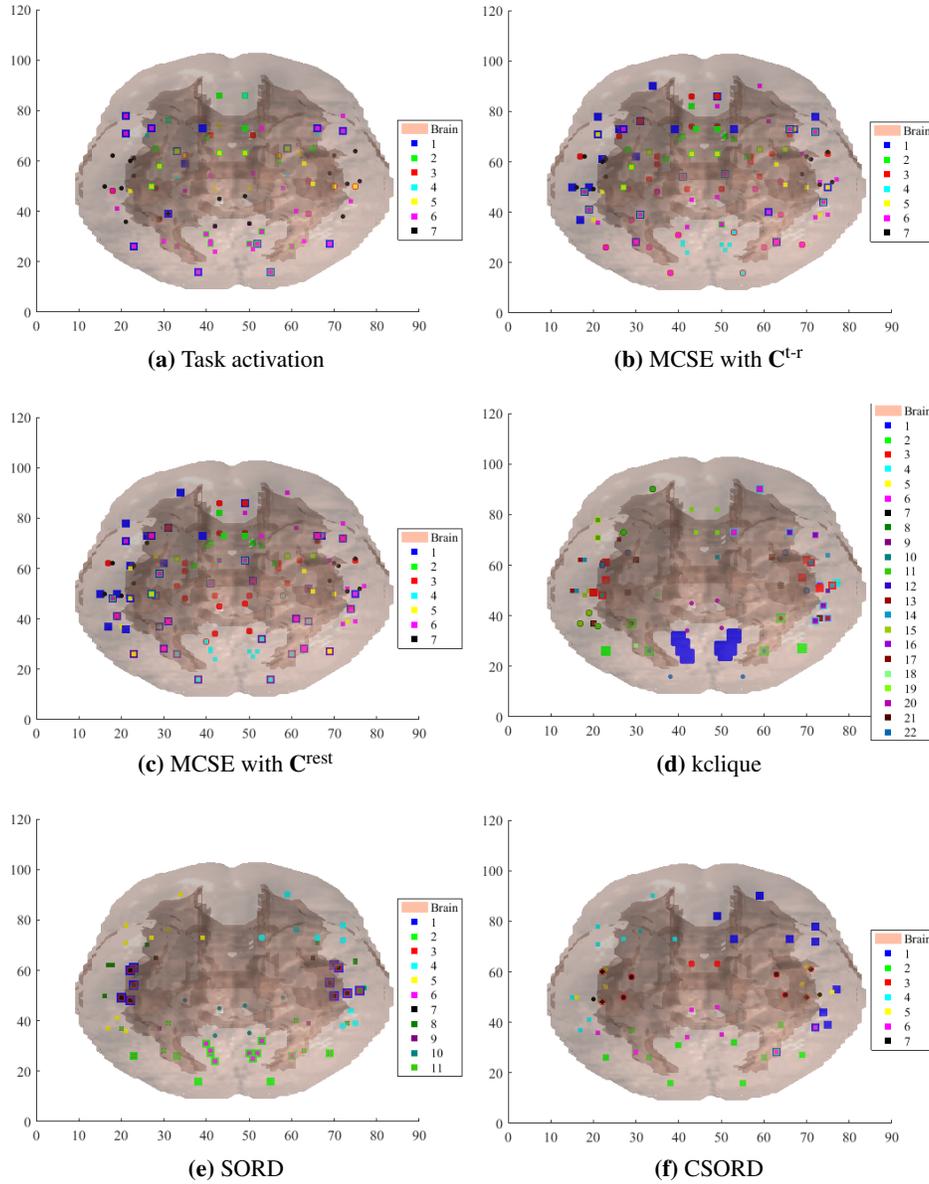


Figure 5.13: Visualization of Task activation and overlapping subnetworks extracted from our proposed approach and contrasted three other methods. The brain is visualized in the axial view. Our proposed MCSE approach outperforms existing state-of-the-art overlapping community detection methods by detecting well-known hubs which reside within subnetwork overlaps.

number of brain regions detected in subnetworks. The ratio of five methods are 0.3482, 0.3482, 0.4444, 0.4328 and 0.2885. Our proposed approach can generate the similar ratio of interacting nodes which reside within subnetwork overlaps to the existing overlapping methods. We note that CSORD generated relatively smaller number of interacting nodes, the possible reason is that the strict stability selection resulted in exclusion of some meaningful nodes, which were taken as false detected nodes arising from noise [203].

By examining the locations of those interacting nodes, we found that our proposed MCSE with C^{t-r} approach identified subnetwork overlaps within pre- and postcentral gyri, medial superior frontal cortex, inferior frontal gyrus, superior parietal lobule, precuneus, lateral occipital cortex, occipital pole and frontal orbital cortex; which match well with functional hubs previously identified by graph-theoretical analysis based on the degree of the voxels [256]. Besides, brain regions of insula, putamen, thalamus, supramarginal gyrus have been found within subnetwork overlaps, which match well with the connector hubs identified using the centrality measures [257]. The results of using MCSE with C^{rest} is very similar to C^{t-r} , only that precuneus cortex was missed, and the temporal pole was misclassified into the subnetwork overlaps. This result confirms the benefit of integrating the information from both task and rest data. Both MCSE methods also identified lingual gyrus and fusiform cortex around as interacting nodes. Lingual gyrus was identified as a hub based on cortical thickness correlation [258] and the fusiform cortex within occipitotemporal cortex has been found to be intermediary “hub” linking visual and higher linguistic representations [259].

As for the traditional kclique approach, biologically meaningful subnetwork overlaps were found within inferior frontal gyrus, superior and middle temporal gyri [256], supramarginal gyrus, insula [257], inferior temporal gyrus [260], and occipitotemporal cortex [259]. kclique failed to identify all the other aforementioned (connector) hubs which were found using our proposed methods. Instead, regions normally were not considered to reside in subnetwork overlaps were found, such as temporal fusiform cortex, central opercular cortex, and parietal operculum cortex. On the other hand, this kclique approach detected angular gyrus (functioning as a semantic hub) within subnetwork overlaps.

SORD was able to find subnetwork overlaps within inferior, superior and mid-

dle temporal gyri, superior parietal lobule, lateral occipital cortex, occipital pole and lingual gyrus that match well with functional hubs, but failed to find other hub regions identified by MCSE. Instead, SORD detected many regions as interacting nodes, which normally are not considered as hubs, such as intracalcarine cortex and cuneal cortex in the visual system, and regions in language related system, including central opercular cortex, parietal operculum cortex, planum polare, planum temporale, heschls gyrus, and supracalcarine cortex.

With relatively lower number of overlapping ratio, CSORD identified biological meaningful subnetwork overlaps within regions such as pre- and postcentral gyri, middle temporal gyrus, angular gyrus and lateral occipital cortex, while failed to find any other hubs. Similar to SORD, it included some regions in language related system to the subnetwork overlaps, such as central opercular cortex, parietal operculum cortex, and planum temporale. We did not discover the single subnetwork constituting the visual corticostriatal loop, striothalamo-cortical loop, and cerebello-thalamo-cortical loop, which was found in [203]. The reason could be this connection was reflected in AC, instead of task functional connectivity.

Collectively, our proposed MCSE approach is able to identify subnetwork overlaps which constitute more biologically meaningful brain regions, such as hubs, compared against contrasted methods.

5.3.3.4 Comparison Between the Subnetwork Overlaps and the RW Posterior Probability

We also examined the overlapping subnetworks derived from our approach MCSE with C^{t-r} by assessing the probability of a node belonging to subnetworks using our own recently proposed multimodal RW approach [6], to verify that our overlapping subnetwork assignments correspond with the posterior probability. The underlying rationale is that for an interacting node, which resides within the subnetwork overlaps, its probability of belonging to a subnetwork will be distributed across the subnetworks it resides in. On the other hand, an individual node, which does not reside within subnetwork overlaps, would have higher chances to possess a dominant probability of being assigned to a particular subnetwork. Hence the difference of probabilities of a node being assigned to the first two subnetworks with the first two highest probabilities indicates the possibility of a node residing within

subnetwork overlaps. Interacting nodes tend to have a smaller value of difference of first two highest probabilities.

We here define the degree of overlapping confidence as the subtraction from one of the difference between the first two highest probabilities of a node being assigned to subnetworks. The nodes identified within the subnetwork overlaps (interacting nodes) and outside of the overlaps (individual nodes) are considered as two populations. For each population, the average overlapping confidence is defined as below in Equation 5.35:

$$overConf = \frac{1}{|S|} \sum_{i \in S} (1 - (p_i^{max} - p_i^{smax})), \quad (5.35)$$

where S is a set of nodes, either nodes residing within or outside the subnetwork overlaps, and p^{max} is the maximal probability of a node belonging to subnetworks, and p^{smax} is the second maximal probability. Thus, the interacting node population is expected to have higher *overConf* compared to individual nodes.

We first derived the probabilities of each node being assigned into all possible subnetworks using our recently proposed multimodal RW approach [6], where two sources of connectivity matrices are \mathbf{C}^{rest} and $\bar{\mathbf{C}}^{task}$, matching with how \mathbf{C}^{t-r} was generated in our approach. The number of seeds within each subnetwork n_k was set to $[2, \dots, 9]$, where 9 is 75% of 12, the minimal number of nodes which are included in non-overlapping subnetwork extraction using \mathbf{C}^{t-r} .

We found that the overlapping confidence of the interacting nodes with an average *overConf* of 0.6884 are statistically higher than the individual nodes with an average *overConf* of 0.6338 based on the Wilcoxon signed rank test at $p=0.006$, see Figure 5.14. This finding confirms that the overlapping subnetwork assignments based on our proposed MCSE match with the probability derived independently from our RW based approach.

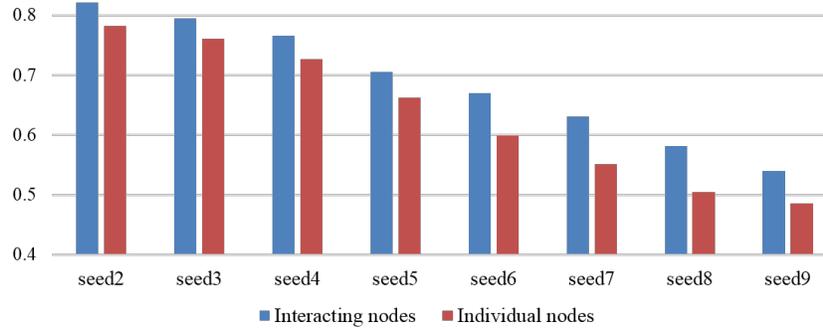


Figure 5.14: Overlapping confidence of interacting nodes in blue versus individual nodes in red derived by MCSE with C^{t-r} . The probability of a node being assigned into subnetworks was derived by the RW based approach [6].

5.3.4 Discussion

5.3.4.1 Benefits of Clique Identification Based on Task Co-activation

The traditional definition of clique is the fully connected subgraphs identified by the connections between brain regions mostly on resting state connectivity. In our approach, we present a novel way to identify cliques based on the similarity of activation patterns between nodes. We argue that the *clique* concept closely resemble the *canonical network components* that are recruited selectively and repeatedly in different task-induced activities [213]. Different from the traditional kclique method [246], our clique-based approach is able to utilize both the task activation information and the task-induced connectivity strength rather than only the binarized connectivity information used in kclique method. Besides, the cliques derived using our approach have flexible clique size, which was determined automatically, having more neuro-scientific justifications than the fixed clique size in [246]. Moreover, we estimate the properties from cliques to indicate the importance of cliques, which gives us a better control over falsely including some fake cliques due to noise. We did find cliques within brain areas that well match with hubs, which indicates that our approach can identify subnetwork overlaps with more biological meaning than the traditional kclique method.

5.3.4.2 Multisource Information Integration Improves the Overlapping Subnetwork Extraction

Compared to the widely used overlapping community detection methods, our approach integrates information from multiple sources. We used both task information (including task activation and connectivity strength) for clique identification and resting state connectivity information for subnetwork membership sharing. The results from reproducibility and biological meaning indicate that our multi-source approach, especially MCSE with C^{t-r} , outperforms uni-source methods such as kclique and SORD, which has been proven to give better overlapping brain subnetwork extraction results compared to state of the art techniques such as OSLOM. We note that our multisource approach further outperformed the multisource version of SORD, CSORD. The reason could be that clique based idea and the sharing of the node subnetwork membership is more straightforward and have more direct biological intuition than relying on survival probabilities of different genders in evolution, which is used in CSORD.

5.3.4.3 Overlapping Subnetwork Identification Corresponds with the RW Posterior Probability

We have identified subnetwork overlaps within brain regions that well match with hubs defined using functional, structural and anatomical information. The results enable us to study the interaction and integration between subnetworks and how interacting nodes (or important hubs) play their roles in the information flow across different subnetworks. We further demonstrated that the assignments of interacting/individual nodes using our proposed MCSE correspond with the posterior probability derived independently from our previously proposed RW based approach [6]. The finding of more distinguishable overlapping confidence between two populations of nodes when the number of seeds was set within a range of [6, 8] confirms the merit of using multiple seeds within a reasonable range (not including connector hubs) in the RW based approach.

5.3.4.4 Other Considerations

We have also discovered that the uni-source traditional kclique approach has high computational complexity when the graph density increases, where there exist large number of fully connected subgraphs. The computational complexity of both SORD and CSORD increases when the bootstrap sampling increases [86]. However, the computation time of our proposed MCSE is quite reasonable and not sensitive to the graph densities.

In terms of the coverage of the brain area from the subnetwork extraction results, SORD and CSORD neglected some brain regions which are not selected as significant nodes by the stability selection. While these two approaches offered this extra feature, they sometimes falsely missed important nodes and failed to cover the whole brain for analysis.

5.3.4.5 Limitations and Future Work

In this work, we presented an approach to identify cliques based on task information and extract overlapping subnetworks using both task and rest data. However, the ideal multimodal framework would be able to integrate AC into the fusion for detecting overlapping subnetworks. The challenge is to discover the relationship between AC and task activation, which enables the clique identification to incorporate anatomical information. Our future work will focus on integrating AC into task-activation based clique identification, or into the multimodal subnetwork membership assignment, e.g., using the multimodal RW approach or the multislice approach [231].

5.4 Summary

In order to further improve brain subnetwork extraction, we explored multimodal/multisource fusion technologies and the higher order relations among network nodes.

First, we extended the RW based subnetwork extraction methods to a multimodal approach for the fusion of AC and FC using provincial hubs to propagate the modular structure information across different modalities. Second, we proposed a high order relation informed model based on hypergraph to integrate both

rest and task information for non-overlapping brain subnetwork extraction. Next, we further proposed multisource overlapping brain subnetwork extraction using canonical network components, i.e., cliques, which we defined based on task co-activation. Based on the clique concept, we investigated overlapping subnetworks based on a label sharing scheme which incorporates the rest data information and task data embedded with higher order relations.

We have demonstrated that integrating multimodal/multisource information and using high order relations result in better subnetwork extraction in terms of the overlaps to well-established brain systems, test-retest repeatability, inter-subject reproducibility and biological meaning.

Chapter 6

Conclusions

In this thesis, we presented our contributions towards improving brain connectivity analysis based on the graph theory representation of the human brain network. We proposed novel multimodal techniques to study brain connectomics through a network-based inquiry into the brain's structure, function and connectivity. We conclude in this chapter the methods we employed within different stages in graphical analysis framework, how they tackled the RQs raised in Section 1.6, and the possible directions we would suggest for future investigations.

6.1 Our Contributions in Graphical Framework for Human Brain Connectivity Analysis

In connectome data graphical representations, human brain connectivity analysis is usually conducted in terms of brain node identification, brain edge building and brain network property analysis. We have made contributions towards each corresponding step in the framework.

6.1.1 Definition of Nodes: Parcellation-based Brain Connectome

Our first contribution is towards improving parcellation, i.e., the brain network node definition. In Chapter 2, we first devised a more reliable method for parcellation (RQ1) by embedding neighborhood connectivity information into the affinity matrix to ameliorate the adverse effects of noise in the neuroimaging data. We were

able to distinguish between voxels at the boundaries and those found in the interior regions of clusters, and to better estimate the affinity values based on the voxels' membership in putative clusters. We have demonstrated that our proposed approach produces parcellations with better intra-subject test-retest reliability, higher inter-subject test-retest reliability and higher subject-group consistency.

In order to investigate whether combining the modalities can actually improve brain parcellation (RQ2), we proposed to integrate the connectivity from both anatomical and functional modalities based on adaptive weighting using voxel-wise test-retest reliability. Our rationale is to overcome the inherent limitations associated with each of the two modalities, by fusing different facets of information from the AC and FC. We further incorporated structural information on the gyri and sulci when performing a regional level extension.

We then designed a number of evaluation metrics to validate our improvements on parcellation (RQ3). We quantitatively demonstrated the superiority of multimodal parcellation as compared with unimodal parcellation, and existing, widely used functional and anatomical atlases in terms of reproducibility, functional homogeneity, and leftout data likelihood. Multimodal parcellations also overlap with known cytoarchitectonic areas, and the extracted subnetworks matched well with the established brain systems.

Our results demonstrated that incorporating neighborhood information and integrating multiple brain attributes that reflect different aspects of the brain's organization do indeed improve brain parcellation.

6.1.2 Definition of Edges: Estimating Connections Between Brain Regions

In Chapter 3, we proposed the employment of noise reduction techniques in the brain edge building step. The edge weights are derived from the connectivity estimates, which depend greatly on the quality of fMRI or dMRI data. Both modalities suffer from image resolution limitations and pronounced noise, which might obscure the brain network topology.

First, to combat false negatives in connectivity estimation (RQ4), we proposed a matrix completion based technique, which (i) formulates the recovering missing connection problem as a low rank matrix completion problem, (ii) tackles the non-

trivial rank estimation using the information aggregation approach, and (iii) exploits neighborhood information to solve the negative entry problem. Based on synthetic data and real HCP data, we quantitatively demonstrated the superiority of our approach as compared with the existing state-of-the-art methods in terms of recovery accuracy and IQ prediction.

Next, we proposed a local thresholding method to suppress the false positive connectivity estimates (RQ5), which tackled the regional bias problem in connectivity estimation. We showed that our local thresholding method achieved better performance results in brain subnetwork extraction application in terms of its accuracy and reproducibility.

6.1.3 Network Measures: Graphical Metrics for Brain Connectomics

Our third contribution was to improve brain network property analyses, particularly the subnetwork extraction, by using multi-pronged graphical metric guided methods. In Chapter 4, we investigated whether incorporating more domain-related graphical metrics improves subnetwork extractions (RQ6), and explored if a model that resembles the brain subnetwork's biological nature improves subnetwork extractions (RQ7). Specifically, we first proposed a modularity reinforcement model based on connection fingerprints which highlight the putative modular structure of brain graphs. The underlying assumption is that node pairs belonging to the same subnetwork presumably connect to a similar set of brain regions, i.e., have similar connection fingerprints. By applying modularity reinforcement, we demonstrated that our approach attained higher accuracy in subnetwork extraction as compared to conventional community detection methods on synthetic data. A higher degree of overlap with established brain systems and a higher subnetwork reproducibility were also shown in the real data.

In order to incorporate a greater number of graphical metrics, we have further presented a random walker based provincial hub guided model, which utilizes informative, module-related network metrics including hubs, within-module degree scores, and participant coefficients. At the same time, we argued that the manner in which the nodes are clustered into groups, based on the probabilities of walking to seeds in the Random Walker, closely resembles the mechanism whereby

brain regions within a subnetwork are inter-linked via provincial hubs. We have further devised a multi-seed strategy to reinforce robustness against noisy connectivity connections. The results on synthetic data from the subnetwork extraction accuracy and the overlaps to well-known brain systems in the real data have both demonstrated the clear benefits gained in applying multi-pronged graphical metrics and devising a model based on biological intuition for subnetwork extraction.

6.1.4 Multimodal/Multisource Fusion to Improve Brain Connectivity Analysis

Our fourth contribution was to develop multimodal/multisource fusion techniques to further improve brain subnetwork extraction. In Chapter 5, we investigated the feasibility and benefits of applying the multimodal fusion (RQ8) and multi-source integration approaches towards studying higher order relations among network nodes (RQ9).

We first demonstrated that anatomical and functional data can be fused together by the provincial hubs to propagate modular structural information across different modalities. Our proposed random walker based approach facilitates multimodal integration by updating the augmented prior edge linking nodes to those provincial hubs connecting the anatomical and functional subnetworks. Moreover, we were able to infer the probability of a node being assigned to a subnetwork for a further investigation of the overlapping subnetworks. We showed that our multimodal approach achieved a significantly higher subnetwork identification accuracy than a number of state-of-the-art approaches on synthetic data. We also demonstrated that our estimated subnetworks matched well with established brain systems and attained comparable or higher inter-subject reproducibility when applied to real data.

Secondly, we proposed a high order relation informed approach based on hypergraph to combine the information from multi-task data and resting state data to improve subnetwork extraction. Our assumption is that task data can be beneficial for the subnetwork extraction process, since the repeatedly activated nodes involved in diverse tasks might be the canonical network components which comprise pre-existing repertoires of resting state subnetworks [213]. Our proposed high order relation informed subnetwork extraction are based on a strength information

embedded hypergraph. Our approach (1) facilitates the multisource integration for subnetwork extraction, (2) utilizes information on relationships and changes between the nodes across different tasks, and (3) enables the study on higher order relations among brain network nodes. We demonstrated that fusing task activation, task-induced connectivity and resting state functional connectivity based on hypergraphs improves subnetwork extraction compared to employing a single source from either rest or task data in terms of subnetwork modularity measure, inter-subject reproducibility, along with more biologically meaningful subnetwork assignments.

We further studied the overlapping brain subnetwork extraction using cliques, which we defined as co-activated node groups performing multiple tasks. We proposed a multisource subnetwork extraction approach based on the co-activated *clique*, which (1) uses task co-activation and task connectivity strength information for clique identification, (2) automatically detects cliques of different sizes having more neuroscientific justifications, and (3) shares the subnetwork membership, derived from a fusion of rest and task data, among the nodes within a clique for overlapping subnetwork extraction. Compared to the commonly used overlapping community detection techniques, we showed that our approach improved subnetwork extraction in terms of group-level and subject-wise reproducibility. We also showed that our multisource approach identified subnetwork overlaps within brain regions that matched well with hubs defined using functional and anatomical information, which enables us to study the interactions between the subnetworks and how hubs play their role in information flow across different subnetworks. We further demonstrated that the assignments of interacting/individual nodes using our approach correspond with the posterior probability derived independently from our multimodal random walker based approach.

Our results collectively indicate that the multimodal fusion of resting FC and AC information and the multisource integration of task and rest functional data outperform classical unimodal brain connectivity analysis methods.

6.2 Future Work

In this work, we mostly focus on studying human brain connectivity from multiple sources of anatomical and functional connectivity using graph theory. In Chapter 2, we fused the AC and FC for improved parcellation based on an adaptive weighting scheme. We have proved that our approach provides improved parcellation in practice; however, understanding the exact nature of the relationship between AC and FC warrants further research. With the acquisition of a greater degree of knowledge on this subject, studies on fusing the two modalities based on advanced theories, and/or applying deep learning methods and reformulating parcellation as a voxel-wise classification problem might give us new insight into multimodal parcellation.

In Chapter 3, we focused on denoising connectivity strength which is used as graph edge weights. Future improvements in the quality of data collection with advanced MRI scans and more reasonable protocols will greatly benefit the accuracy of connectivity estimation. Throughout the thesis, we used the most commonly employed Pearson's correlation coefficient between regional activity time courses. Alternative measures such as partial correlations, mutual information and coherence might lead to some new discoveries or confirm our current findings. Also, our work employed the AC and FC to study undirected brain graphs. Another promising direction is to incorporate the effective connectivity when estimating graph edges using functional data with a higher temporal resolution, such as EEG data, which enables more accurate EC estimations using spectral coherence or Granger causality measures [75]. This direction will include one more aspect for multimodal fusion which could be based on multi-layer graph, while simultaneously allows for studies on directed graphs.

In Chapter 4 and Chapter 5, we made contributions towards improved brain subnetwork extraction by applying multi-pronged graphical metrics and the multimodal/multisource fusion approach. We have linked image and graphical theory analysis tools to biological intuition to closely model the brain's nature. Knowledge of the exact nature of the relationship between all the multimodal/multisource information requires further research. One additional direction would be to consider higher order relations among the network nodes by pursuing further applica-

tions in multilayer graphs in brain connectivity analysis, where dynamic connectivity [261] can be incorporated into the multimodal framework.

In this thesis, we have studied brain connectivity to understand the basic architecture of the healthy human brain. The resulting methods might provide us with important new insights into cognitive and clinical neuroscience, which brings us one step closer to understand the relationships of brain architecture to diseases whose alternations may be described as dysconnectivity syndromes. One promising future research direction is to explore how brain network features can be linked, e.g., to gender, aging, genetic factors, and the stages of disease severity, possibly based on methods such as deep learning approaches.

For example, we think it is feasible to build a deep learning framework for neurological disorder classification or disease stage prediction. Ideally, this model will incorporate both information from the original image data and topological information extracted from the brain connectome, possibly through a multistream strategy. Firstly, the results from our work in this thesis (brain parcellation and edge building) can be used to construct the brain connectome graph for topological information extraction. Secondly, we can design a deep neural network architecture to learn the features by training an unsupervised deep learning scheme based on a big number of healthy controls. Importantly, the results from our current research (such as parcellation, subnetwork extraction, clique identification and hubs) can be utilized to guide the design of the architecture, e.g., using the (hierarchical) connectivity information to design the hidden layers. Incorporating model prior into the framework might improve the performance of deep learning architectures [262, 263], with which it becomes possible to reflect or involve the domain-specific knowledge [124]. Moreover, a good knowledge of the brain connectivity from our work can help explain the features learned from the deep black box, e.g., hidden layers from low to high can be related to attributes derived from nodal features based on parcellation, edge-to-edge relationships, clique-wise node group interaction, subnetwork structure, to hub features or brain network-wide integration features. One other interesting direction is to incorporate clinical prior features [131], such as patient history, age, demographics, behavioural scores, genetic phenotypes and others. Finally, we can fine tune the parameters in the deep learning model using the labeled disease data for specific applications.

We have also considered some directions to handle important steps in the model training. First, the prior knowledge learned from this thesis can be used for weight initialization (e.g., how nodes are closely connected to a subnetwork it belongs to, can be related to how strong the weights connecting units across layers). Secondly, in order to tackle the local minima problem, we suggest to deploy a multimodal perturbation strategy. Thirdly, we suggest to use regularization or ensemble learning (e.g., majority voting from a number of deep network models) to tackle the overfitting problem.

Other directions to improve brain connectivity analysis using machine learning or deep learning approaches include (1) big data collection and label acquisition using methods such as multiple-instance or active learning approaches [131]; (2) designing methods to overcome class imbalance, such as modifying loss function [264]; (3) devising methods suitable for graph representation of brain connectome data [265]; and (4) developing algorithmic techniques such as transfer learning to more efficiently adopt networks trained from the computer vision area and handle cross-institution difference and cross-modality generalization.

Bibliography

- [1] Chendi Wang, Burak Yoldemir, and Rafeef Abugharbieh. Improved functional cortical parcellation using a neighborhood-information-embedded affinity matrix. In *IEEE International Symposium on Biomedical Imaging*, pages 1340–1343. IEEE Press, 2015.
- [2] Chendi Wang, Burak Yoldemir, and Rafeef Abugharbieh. Multimodal cortical parcellation based on anatomical and functional brain connectivity. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 21–28. Springer, 2015.
- [3] Chendi Wang, Bernard Ng, and Rafeef Abugharbieh. Multimodal whole brain parcellation. *submitted*, xx(x):xxxx–xxxx, 2018.
- [4] Chendi Wang, Bernard Ng, Alborz Amir-Khalili, and Rafeef Abugharbieh. Recovering missing connections in diffusion weighted mri using matrix completion. In *International Conference on Medical Image Computing and Computer-Assisted Intervention Workshop on Computational Diffusion MRI*. Springer, 2017.
- [5] Chendi Wang, Bernard Ng, and Rafeef Abugharbieh. Modularity reinforcement for improving brain subnetwork extraction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 132–139. Springer, 2016.
- [6] Chendi Wang, Bernard Ng, and Rafeef Abugharbieh. Multimodal brain subnetwork extraction using provincial hub guided random walks. In *International Conference on Information Processing in Medical Imaging*, pages 287–298. Springer, Cham, 2017.
- [7] Chendi Wang and Rafeef Abugharbieh. Fusion of task and resting state functional connectivity for subnetwork extraction based on hypergraph. *submitted*, xx(x):xxxx–xxxx, 2018.

- [8] Chendi Wang and Rafeef Abugharbieh. Co-activated clique based multisource overlapping subnetwork extraction. *submitted*, xx(x):xxxx–xxxx, 2018.
- [9] GBD 2015 DALYs and HALE Collaborators. Global, regional, and national disability-adjusted life-years (dalys) for 315 diseases and injuries and healthy life expectancy (hale), 1990–2015: a systematic analysis for the global burden of disease study 2015. *Lancet*, 388(October 8):1603–1658, 2016.
- [10] GBD 2015 Disease, Injury Incidence, and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the global burden of disease study 2015. *Lancet*, 388(October 8):1545–1602, 2016.
- [11] Terrence J Sejnowski, Patricia S Churchland, and J Anthony Movshon. Putting big data to good use in neuroscience. *Nature neuroscience*, 17(11):1440–1441, 2014.
- [12] Rolf Kötter. Online retrieval, processing, and visualization of primate connectivity data from the cocomac database. *Neuroinformatics*, 2(2):127–144, 2004.
- [13] Saad Jbabdi and Heidi Johansen-Berg. Tractography: where do we go from here? *Brain connectivity*, 1(3):169–183, 2011.
- [14] R Cameron Craddock, Saad Jbabdi, Chao-Gan Yan, Joshua T Vogelstein, F Xavier Castellanos, Adriana Di Martino, Clare Kelly, Keith Heberlein, Stan Colcombe, and Michael P Milham. Imaging human connectomes at the macroscale. *Nature methods*, 10(6):524–539, 2013.
- [15] Marco Catani. From hodology to function. *Brain*, 130(3):602–605, 2007.
- [16] Alex Fornito, Andrew Zalesky, and Michael Breakspear. Graph analysis of the human connectome: promise, progress, and pitfalls. *Neuroimage*, 80:426–444, 2013.
- [17] Mikail Rubinov and Olaf Sporns. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*, 52(3):1059–1069, 2010.

- [18] Jonathan D Power, Alexander L Cohen, Steven M Nelson, Gagan S Wig, Kelly Anne Barnes, Jessica A Church, Alecia C Vogel, Timothy O Laumann, Fran M Miezin, Bradley L Schlaggar, et al. Functional network organization of the human brain. *Neuron*, 72(4):665–678, 2011.
- [19] Nicolas A Crossley, Andrea Mechelli, Petra E Vértes, Toby T Winton-Brown, Ameera X Patel, Cedric E Ginestet, Philip McGuire, and Edward T Bullmore. Cognitive relevance of the community structure of the human brain functional coactivation network. *Proceedings of the National Academy of Sciences*, 110(28):11583–11588, 2013.
- [20] David C Van Essen, Matthew F Glasser, Donna L Dierker, John Harwell, and Timothy Coalson. Parcellations and hemispheric asymmetries of human cerebral cortex analyzed on surface-based atlases. *Cerebral cortex*, 22(10):2241–2262, 2011.
- [21] Nikolaus Weiskopf, Chloe Hutton, Oliver Josephs, Robert Turner, and Ralf Deichmann. Optimized epi for fmri studies of the orbitofrontal cortex: compensation of susceptibility-induced gradients in the readout direction. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 20(1):39–49, 2007.
- [22] Aaron F Alexander-Bloch, Nitin Gogtay, David Meunier, Rasmus Birn, Liv Clasen, Francois Lalonde, Rhoshel Lenroot, Jay Giedd, and Edward T Bullmore. Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia. *Frontiers in systems neuroscience*, 4:147, 2010.
- [23] WR Shirer, S Ryali, E Rykhlevskaia, V Menon, and MD Greicius. Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cerebral cortex*, 22(1):158–165, 2012.
- [24] Public Health Agency of Canada. Mapping connections: An understanding of neurological conditions in canada. *The National Population Health Study of Neurological Conditions*, 2014.
- [25] Cornelis J Stam. Modern network science of neurological disorders. *Nature Reviews Neuroscience*, 15(10):683–695, 2014.
- [26] Olaf Sporns, Giulio Tononi, and Rolf Kötter. The human connectome: a structural description of the human brain. *PLoS computational biology*, 1(4):e42, 2005.

- [27] MJ Rosa, J Daunizeau, and KJ Friston. Eeg-fmri integration: a critical review of biophysical modeling and data analysis approaches. *Journal of integrative neuroscience*, 9(04):453–476, 2010.
- [28] Michael E Phelps. *PET: molecular imaging and its biological applications*. Springer Science & Business Media, 2004.
- [29] Seiji Ogawa, Tso-Ming Lee, Alan R Kay, and David W Tank. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, 87(24):9868–9872, 1990.
- [30] Scott A Huettel, Allen W Song, and Gregory McCarthy. *Functional magnetic resonance imaging*, volume 1. Sinauer Associates Sunderland, 2004.
- [31] Robert Turner and Terry Jones. Techniques for imaging neuroscience. *British medical bulletin*, 65(1):3–20, 2003.
- [32] Michael C Stevens. The contributions of resting state and task-based functional connectivity studies to our understanding of adolescent brain network maturation. *Neuroscience & Biobehavioral Reviews*, 70:13–32, 2016.
- [33] Yong He, Zhang J Chen, and Alan C Evans. Small-world anatomical networks in the human brain revealed by cortical thickness from mri. *Cerebral cortex*, 17(10):2407–2419, 2007.
- [34] Peter J Basser, James Mattiello, and Denis LeBihan. Estimation of the effective self-diffusion tensor from the nmr spin echo. *Journal of Magnetic Resonance, Series B*, 103(3):247–254, 1994.
- [35] Haz-Edine Assemlal, David Tschumperlé, Luc Brun, and Kaleem Siddiqi. Recent advances in diffusion mri modeling: Angular and radial reconstruction. *Medical image analysis*, 15(4):369–396, 2011.
- [36] Martijn P Van Den Heuvel and Hilleke E Hulshoff Pol. Exploring the brain network: a review on resting-state fmri functional connectivity. *European Neuropsychopharmacology*, 20(8):519–534, 2010.
- [37] Tianzi Jiang. Brainnetome: a new-ome to understand the brain and its disorders. *Neuroimage*, 80:263–272, 2013.

- [38] Marwan N Baliki, Paul Y Geha, A Vania Apkarian, and Dante R Chialvo. Beyond feeling: chronic pain hurts the brain, disrupting the default-mode network dynamics. *The Journal of Neuroscience*, 28(6):1398–1403, 2008.
- [39] Zhenyu Zhou, Yonghong Chen, Mingzhou Ding, Paul Wright, Zuhong Lu, and Yijun Liu. Analyzing brain networks with pca and conditional granger causality. *Human brain mapping*, 30(7):2197–2206, 2009.
- [40] M De Luca, CF Beckmann, N De Stefano, PM Matthews, and Stephen M Smith. fmri resting state networks define distinct modes of long-distance interactions in the human brain. *Neuroimage*, 29(4):1359–1367, 2006.
- [41] Chris HQ Ding, Xiaofeng He, and Horst D Simon. On the equivalence of nonnegative matrix factorization and spectral clustering. In *SDM*, volume 5, pages 606–610. SIAM, 2005.
- [42] Franco Cauda, Federico D’Agata, Katiuscia Sacco, Sergio Duca, Giuliano Geminiani, and Alessandro Vercelli. Functional connectivity of the insula in the resting brain. *Neuroimage*, 55(1):8–23, 2011.
- [43] Raymond Salvador, John Suckling, Christian Schwarzbauer, and Ed Bullmore. Undirected graphs of frequency-dependent functional connectivity in whole brain networks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1457):937–946, 2005.
- [44] Bertrand Thirion, Silke Dodel, and Jean-Baptiste Poline. Detection of signal synchronizations in resting-state fmri datasets. *Neuroimage*, 29(1):321–327, 2006.
- [45] CJ Stam, W De Haan, A Daffertshofer, BF Jones, I Manshanden, AM Van Cappellen Van Walsum, T Montez, JPA Verbunt, JC De Munck, BW Van Dijk, et al. Graph theoretical analysis of magnetoencephalographic functional connectivity in alzheimer’s disease. *Brain*, 132(1):213–224, 2009.
- [46] Danielle S Bassett and Olaf Sporns. Network neuroscience. *Nature neuroscience*, 20(3):353, 2017.
- [47] Alex Fornito and Edward T Bullmore. Connectomics: a new paradigm for understanding brain disease. *European Neuropsychopharmacology*, 25(5):733–748, 2015.

- [48] JG White, E Southgate, JN Thomson, and S Brenner. The structure of the nervous system of the nematode *caenorhabditis elegans*: the mind of a worm. *Phil. Trans. R. Soc. Lond.*, 314:1–340, 1986.
- [49] Ravi D Mill, Takuya Ito, and Michael W Cole. From connectome to cognition: The search for mechanism in human functional brain networks. *NeuroImage*, 2017.
- [50] Xi-Nian Zuo, Ye He, Richard F Betzel, Stan Colcombe, Olaf Sporns, and Michael P Milham. Human connectomics across the life span. *Trends in cognitive sciences*, 21(1):32–45, 2017.
- [51] NA Crossley, PT Fox, and ET Bullmore. Meta-connectomics: human brain network and connectivity meta-analyses. *Psychological medicine*, 46(5):897–907, 2016.
- [52] Cornelis J Stam. Modern network science of neurological disorders. *Nature Reviews Neuroscience*, 15(10):683–695, 2014.
- [53] Andrew Zalesky, Alex Fornito, Ian H Harding, Luca Cocchi, Murat Yücel, Christos Pantelis, and Edward T Bullmore. Whole-brain anatomical networks: does the choice of nodes matter? *Neuroimage*, 50(3):970–983, 2010.
- [54] Matthew F Glasser, Timothy S Coalson, Emma C Robinson, Carl D Hacker, John Harwell, Essa Yacoub, Kamil Ugurbil, Jesper Andersson, Christian F Beckmann, Mark Jenkinson, et al. A multi-modal parcellation of human cerebral cortex. *Nature*, 536(7615):171–178, 2016.
- [55] Matthew L Stanley, Malaak N Moussa, Brielle M Paolini, Robert G Lyday, Jonathan H Burdette, and Paul J Laurienti. Defining nodes in complex brain networks. *Frontiers in computational neuroscience*, 7, 2013.
- [56] Marcel A de Reus and Martijn P Van den Heuvel. The parcellation-based connectome: limitations and extensions. *Neuroimage*, 80:397–404, 2013.
- [57] Rahul S Desikan, Florent Ségonne, Bruce Fischl, Brian T Quinn, Bradford C Dickerson, Deborah Blacker, Randy L Buckner, Anders M Dale, R Paul Maguire, Bradley T Hyman, et al. An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. *Neuroimage*, 31(3):968–980, 2006.
- [58] Laurence J Garey. *Brodmann's 'localisation in the cerebral cortex'*. World Scientific, 1994.

- [59] Matthew F Glasser and David C Van Essen. Mapping human cortical areas in vivo based on myelin content as revealed by t1-and t2-weighted mri. *Journal of Neuroscience*, 31(32):11597–11616, 2011.
- [60] Svenja Caspers, Axel Schleicher, Mareike Bacha-Trams, Nicola Palomero-Gallagher, Katrin Amunts, and Karl Zilles. Organization of the human inferior parietal lobule based on receptor architectonics. *Cerebral Cortex*, 23(3):615–628, 2012.
- [61] Bruce Fischl, Niranjini Rajendran, Evelina Busa, Jean Augustinack, Oliver Hinds, BT Thomas Yeo, Hartmut Mohlberg, Katrin Amunts, and Karl Zilles. Cortical folding patterns and predicting cytoarchitecture. *Cerebral cortex*, 18(8):1973–1980, 2007.
- [62] Brian A Wandell, Serge O Dumoulin, and Alyssa A Brewer. Visual field maps in human cortex. *Neuron*, 56(2):366–383, 2007.
- [63] Feng Shi, Yong Liu, Tianzi Jiang, Yuan Zhou, Wanlin Zhu, Jiefeng Jiang, Haihong Liu, and Zhening Liu. Regional homogeneity and anatomical parcellation for fmri image classification: application to schizophrenia and normal controls. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*, pages 136–143, 2007.
- [64] Yong Yang, Lingzhong Fan, Congying Chu, Junjie Zhuo, Jiaojian Wang, Peter T Fox, Simon B Eickhoff, and Tianzi Jiang. Identifying functional subdivisions in the human brain using meta-analytic activation modeling-based parcellation. *Neuroimage*, 124:300–309, 2016.
- [65] Heidi Johansen-Berg, TEJ Behrens, MD Robson, I Drobnjak, MFS Rushworth, JM Brady, SM Smith, DJ Higham, and PM Matthews. Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 101(36):13335–13340, 2004.
- [66] Sandrine Lefranc, Pauline Roca, Matthieu Perrot, Cyril Poupon, Denis Le Bihan, Jean-François Mangin, and Denis Rivière. Groupwise connectivity-based parcellation of the whole human cortical surface using watershed-driven dimension reduction. *Medical image analysis*, 30:11–29, 2016.
- [67] Guillermo Gallardo, William Wells, Rachid Deriche, and Demian Wassermann. Groupwise structural parcellation of the whole cortex: a logistic random effects model based approach. *NeuroImage*, 2017.

- [68] Gaël Varoquaux, Alexandre Gramfort, Fabian Pedregosa, Vincent Michel, and Bertrand Thirion. Multi-subject dictionary learning to segment an atlas of brain spontaneous activity. In *Biennial International Conference on Information Processing in Medical Imaging*, pages 562–573. Springer, 2011.
- [69] BT Thomas Yeo, Fenna M Krienen, Jorge Sepulcre, Mert R Sabuncu, Danial Lashkari, Marisa Hollinshead, Joshua L Roffman, Jordan W Smoller, Lilla Zöllei, Jonathan R Polimeni, et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of neurophysiology*, 106(3):1125–1165, 2011.
- [70] David C Van Essen and Kamil Ugurbil. The future of the human connectome. *Neuroimage*, 62(2):1299–1310, 2012.
- [71] Alan Tucholka, Bertrand Thirion, Matthieu Perrot, Philippe Pinel, Jean-François Mangin, and Jean-Baptiste Poline. Probabilistic anatomo-functional parcellation of the cortex: how many regions? In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 399–406. Springer, 2008.
- [72] Viktor Jirsa, Olaf Sporns, Michael Breakspear, Gustavo Deco, and Anthony Randal McIntosh. Towards the virtual brain: network modeling of the intact and the damaged brain. *Archives italiennes de biologie*, 148(3):189–205, 2010.
- [73] Annette M Molinaro, Richard Simon, and Ruth M Pfeiffer. Prediction error estimation: a comparison of resampling methods. *Bioinformatics*, 21(15):3301–3307, 2005.
- [74] J-F Mangin, Pierre Fillard, Yann Cointepas, Denis Le Bihan, Vincent Frouin, and Cyril Poupon. Toward global tractography. *NeuroImage*, 80:290–296, 2013.
- [75] Linda J Larson-Prior, Robert Oostenveld, Stefania Della Penna, G Michalareas, F Prior, Abbas Babajani-Feremi, J-M Schoffelen, Laura Marzetti, Francesco de Pasquale, F Di Pompeo, et al. Adding dynamics to the human connectome project with meg. *Neuroimage*, 80:190–201, 2013.
- [76] Peter J Basser, Sinisa Pajevic, Carlo Pierpaoli, Jeffrey Duda, and Akram Aldroubi. In vivo fiber tractography using dt-mri data. *Magnetic resonance in medicine*, 44(4):625–632, 2000.

- [77] David S Tuch. Q-ball imaging. *Magnetic resonance in medicine*, 52(6):1358–1372, 2004.
- [78] Iman Aganj, Christophe Lenglet, Guillermo Sapiro, Essa Yacoub, Kamil Ugurbil, and Noam Harel. Reconstruction of the orientation distribution function in single-and multiple-shell q-ball imaging within constant solid angle. *Magnetic Resonance in Medicine*, 64(2):554–566, 2010.
- [79] Timothy EJ Behrens, H Johansen Berg, Saad Jbabdi, Matthew FS Rushworth, and Mark W Woolrich. Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage*, 34(1):144–155, 2007.
- [80] Marco Reisert, Irina Mader, Constantin Anastasopoulos, Matthias Weigel, Susanne Schnell, and Valerij Kiselev. Global fiber reconstruction becomes practical. *Neuroimage*, 54(2):955–962, 2011.
- [81] BW Kreher, I Mader, and VG Kiselev. Gibbs tracking: a novel approach for the reconstruction of neuronal pathways. *Magnetic Resonance in Medicine*, 60(4):953–963, 2008.
- [82] Thorsten Kahnt, Luke J Chang, Soyoung Q Park, Jakob Heinzle, and John-Dylan Haynes. Connectivity-based parcellation of the human orbitofrontal cortex. *Journal of Neuroscience*, 32(18):6240–6250, 2012.
- [83] Rachid Deriche. Computational brain connectivity mapping: A core health and scientific challenge, 2016.
- [84] Jun Yi Wang, Hervé Abdi, Khamid Bakhadirov, Ramon Diaz-Arrastia, and Michael D Devous. A comprehensive reliability assessment of quantitative diffusion tensor tractography. *Neuroimage*, 60(2):1127–1138, 2012.
- [85] Hanbo Chen, Kaiming Li, Dajiang Zhu, Xi Jiang, Yixuan Yuan, Peili Lv, Tuo Zhang, Lei Guo, Dinggang Shen, and Tianming Liu. Inferring group-wise consistent multimodal brain networks via multi-view spectral clustering. *IEEE transactions on medical imaging*, 32(9):1576–1586, 2013.
- [86] Burak Yoldemir. *Multimodal fusion for assessing functional segregation and integration in the human brain*. PhD thesis, University of British Columbia, 2016.
- [87] Andrew L Alexander, Khader M Hasan, Mariana Lazar, Jay S Tsuruda, and Dennis L Parker. Analysis of partial volume effects in diffusion-tensor mri. *Magnetic Resonance in Medicine*, 45(5):770–780, 2001.

- [88] Stephen M Smith, Karla L Miller, Gholamreza Salimi-Khorshidi, Matthew Webster, Christian F Beckmann, Thomas E Nichols, Joseph D Ramsey, and Mark W Woolrich. Network modelling methods for fmri. *Neuroimage*, 54(2):875–891, 2011.
- [89] Mark Fiecas, Hernando Ombao, Dan van Lunen, Richard Baumgartner, Alexandre Coimbra, and Dai Feng. Quantifying temporal correlations: a test–retest evaluation of functional connectivity in resting-state fmri. *NeuroImage*, 65:231–241, 2013.
- [90] Kevin Murphy, Rasmus M Birn, and Peter A Bandettini. Resting-state fmri confounds and cleanup. *Neuroimage*, 80:349–359, 2013.
- [91] Gael Varoquaux and Bertrand Thirion. How machine learning is shaping cognitive neuroimaging. *GigaScience*, 3(1):28, 2014.
- [92] Mikail Rubinov and Olaf Sporns. Weight-conserving characterization of complex functional brain networks. *Neuroimage*, 56(4):2068–2079, 2011.
- [93] Pawel Skudlarski, Kanchana Jagannathan, Vince D Calhoun, Michelle Hampson, Beata A Skudlarska, and Godfrey Pearlson. Measuring brain connectivity: diffusion tensor imaging validates resting state temporal correlations. *Neuroimage*, 43(3):554–561, 2008.
- [94] Heidi Johansen-Berg. Human connectomics-what will the future demand? *NeuroImage*, 80:541–544, 2013.
- [95] Giulio Tononi, Olaf Sporns, and Gerald M Edelman. A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proceedings of the National Academy of Sciences*, 91(11):5033–5037, 1994.
- [96] Michelle Girvan and Mark EJ Newman. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12):7821–7826, 2002.
- [97] Mark EJ Newman. Fast algorithm for detecting community structure in networks. *Physical review E*, 69(6):066133, 2004.
- [98] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008, 2008.

- [99] Luca Ferrarini, Ilya M Veer, Evelinda Baerends, Marie-José van Tol, Remco J Renken, Nic JA van der Wee, Dirk Veltman, André Aleman, Frans G Zitman, Brenda WJH Penninx, et al. Hierarchical functional modularity in the resting-state human brain. *Human brain mapping*, 30(7):2220–2231, 2009.
- [100] Santo Fortunato and Marc Barthélemy. Resolution limit in community detection. *Proceedings of the National Academy of Sciences*, 104(1):36–41, 2007.
- [101] Kaiming Li, Lei Guo, Jingxin Nie, Gang Li, and Tianming Liu. Review of methods for functional brain connectivity detection using fmri. *Computerized Medical Imaging and Graphics*, 33(2):131–139, 2009.
- [102] Anthony R McIntosh and Bratislav Mišić. Multivariate statistical analyses for neuroimaging data. *Annual review of psychology*, 64:499–525, 2013.
- [103] Martin J McKeown, Scott Makeig, Greg G Brown, Tzyy-Ping Jung, Sandra S Kindermann, Anthony J Bell, and Terrence J Sejnowski. Analysis of fmri data by blind separation into independent spatial components. Technical report, NAVAL HEALTH RESEARCH CENTER SAN DIEGO CA, 1997.
- [104] VA Traag and Jeroen Bruggeman. Community detection in networks with positive and negative links. *Physical Review E*, 80(3):036115, 2009.
- [105] Patric Hagmann, Leila Cammoun, Xavier Gigandet, Reto Meuli, Christopher J Honey, Van J Wedeen, and Olaf Sporns. Mapping the structural core of human cerebral cortex. *PLoS biology*, 6(7):e159, 2008.
- [106] Olaf Sporns and Richard F Betzel. Modular brain networks. *Annual review of psychology*, 67:613–640, 2016.
- [107] Richard F Betzel, Alessandra Griffa, Andrea Avena-Koenigsberger, Joaquín Goñi, Jean-Philippe Thiran, Patric Hagmann, and Olaf Sporns. Multi-scale community organization of the human structural connectome and its relationship with resting-state functional connectivity. *Network Science*, 1(3):353–373, 2013.
- [108] Christian Lohse, Danielle S Bassett, Kelvin O Lim, and Jean M Carlson. Resolving anatomical and functional structure in human brain organization: identifying mesoscale organization in weighted network representations. *PLoS computational biology*, 10(10):e1003712, 2014.

- [109] Timothy O Laumann, Evan M Gordon, Babatunde Adeyemo, Abraham Z Snyder, Sung Jun Joo, Mei-Yen Chen, Adrian W Gilmore, Kathleen B McDermott, Steven M Nelson, Nico UF Dosenbach, et al. Functional system and areal organization of a highly sampled individual human brain. *Neuron*, 87(3):657–670, 2015.
- [110] Angela R Laird, Jack J Lancaster, and Peter T Fox. Brainmap. *Neuroinformatics*, 3(1):65–77, 2005.
- [111] Stephen M Smith, Peter T Fox, Karla L Miller, David C Glahn, P Mickle Fox, Clare E Mackay, Nicola Filippini, Kate E Watkins, Roberto Toro, Angela R Laird, et al. Correspondence of the brain’s functional architecture during activation and rest. *Proceedings of the National Academy of Sciences*, 106(31):13040–13045, 2009.
- [112] I Tavor, O Parker Jones, RB Mars, SM Smith, TE Behrens, and S Jbabdi. Task-free mri predicts individual differences in brain activity during task performance. *Science*, 352(6282):216–220, 2016.
- [113] Jessica S Damoiseaux and Michael D Greicius. Greater than the sum of its parts: a review of studies combining structural connectivity and resting-state functional connectivity. *Brain Structure and Function*, 213(6):525–533, 2009.
- [114] Christopher J Honey, O Sporns, Leila Cammoun, Xavier Gigandet, Jean-Philippe Thiran, Reto Meuli, and Patric Hagmann. Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences*, 106(6):2035–2040, 2009.
- [115] Jing Sui, Rene Huster, Qingbao Yu, Judith M Segall, and Vince D Calhoun. Function–structure associations of the brain: evidence from multimodal connectivity and covariance studies. *Neuroimage*, 102:11–23, 2014.
- [116] Christopher J Honey, Jean-Philippe Thivierge, and Olaf Sporns. Can structure predict function in the human brain? *Neuroimage*, 52(3):766–776, 2010.
- [117] Olaf Sporns. The human connectome: origins and challenges. *Neuroimage*, 80:53–61, 2013.
- [118] Archana Venkataraman, Yogesh Rathi, Marek Kubicki, Carl-Fredrik Westin, and Polina Golland. Joint modeling of anatomical and functional

connectivity for population studies. *Medical Imaging, IEEE Transactions on*, 31(2):164–182, 2012.

- [119] F DuBois Bowman, Lijun Zhang, Gordana Derado, and Shuo Chen. Determining functional connectivity using fmri data with diffusion-based anatomical weighting. *NeuroImage*, 62(3):1769–1779, 2012.
- [120] Farras Abdelnour, Henning U Voss, and Ashish Raj. Network diffusion accurately models the relationship between structural and functional brain connectivity networks. *Neuroimage*, 90:335–347, 2014.
- [121] Lingzhong Fan, Hai Li, Junjie Zhuo, Yu Zhang, Jiaojian Wang, Liangfu Chen, Zhengyi Yang, Congying Chu, Sangma Xie, Angela R Laird, et al. The human brainnetome atlas: a new brain atlas based on connectional architecture. *Cerebral Cortex*, 26(8):3508–3526, 2016.
- [122] Dajiang Zhu, Tuo Zhang, Xi Jiang, Xintao Hu, Hanbo Chen, Ning Yang, Jinglei Lv, Junwei Han, Lei Guo, and Tianming Liu. Fusing dti and fmri data: a survey of methods and applications. *NeuroImage*, 102:184–191, 2014.
- [123] Vince D Calhoun and Jing Sui. Multimodal fusion of brain imaging data: A key to finding the missing link (s) in complex mental illness. *Biological psychiatry: cognitive neuroscience and neuroimaging*, 1(3):230–244, 2016.
- [124] Dinggang Shen, Guorong Wu, and Heung-Il Suk. Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 2017.
- [125] Youngjin Yoo, Tom Brosch, Anthony Traboulsee, David KB Li, and Roger Tam. Deep learning of image features from unlabeled data for multiple sclerosis lesion segmentation. In *International Workshop on Machine Learning in Medical Imaging*, pages 117–124. Springer, 2014.
- [126] Pavel Dvorak and Bjoern H Menze. Local structure prediction with convolutional neural networks for multimodal brain tumor segmentation. In *MCV@ MICCAI*, pages 59–71, 2015.
- [127] Shaoyu Wang, Minjeong Kim, Guorong Wu, and Dinggang Shen. Scalable high performance image registration framework by unsupervised deep feature representations learning. In *Deep Learning for Medical Image Analysis*, pages 245–269. Elsevier, 2017.

- [128] Siqi Liu, Sidong Liu, Weidong Cai, Sonia Pujol, Ron Kikinis, and Dagan Feng. Early diagnosis of alzheimer’s disease with deep learning. In *Biomedical Imaging (ISBI), 2014 IEEE 11th International Symposium on*, pages 1015–1018. IEEE, 2014.
- [129] Heung-Il Suk, Seong-Whan Lee, Dinggang Shen, Alzheimer’s Disease Neuroimaging Initiative, et al. Latent feature representation with stacked auto-encoder for ad/mci diagnosis. *Brain Structure and Function*, 220(2):841–859, 2015.
- [130] Jun Shi, Xiao Zheng, Yan Li, Qi Zhang, and Shihui Ying. Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of alzheimer’s disease. *IEEE journal of biomedical and health informatics*, 2017.
- [131] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghahfoorian, Jeroen AWM van der Laak, Bram van Ginneken, and Clara I Sánchez. A survey on deep learning in medical image analysis. *arXiv preprint arXiv:1702.05747*, 2017.
- [132] Mikael Henaff, Joan Bruna, and Yann LeCun. Deep convolutional networks on graph-structured data. *arXiv preprint arXiv:1506.05163*, 2015.
- [133] Jeremy Kawahara, Colin J Brown, Steven P Miller, Brian G Booth, Vann Chau, Ruth E Grunau, Jill G Zwicker, and Ghassan Hamarneh. Brainnetcn: Convolutional neural networks for brain networks; towards predicting neurodevelopment. *NeuroImage*, 146:1038–1049, 2017.
- [134] Sarah Parisot, Sofia Ira Ktena, Enzo Ferrante, Matthew Lee, Ricardo Guerrero Moreno, Ben Glocker, and Daniel Rueckert. Spectral graph convolutions on population graphs for disease prediction. *arXiv preprint arXiv:1703.03020*, 2017.
- [135] David I Shuman, Benjamin Ricaud, and Pierre Vandergheynst. A windowed graph fourier transform. In *Statistical Signal Processing Workshop (SSP), 2012 IEEE*, pages 133–136. Ieee, 2012.
- [136] Wenqi Li, Guotai Wang, Lucas Fidon, Sebastien Ourselin, M Jorge Cardoso, and Tom Vercauteren. On the compactness, efficiency, and representation of 3d convolutional networks: Brain parcellation as a pretext task. In *International Conference on Information Processing in Medical Imaging*, pages 348–360. Springer, 2017.

- [137] Hannah Spitzer, Katrin Amunts, Stefan Harmeling, and Timo Dickscheid. Parcellation of visual cortex on high-resolution histological brain sections using convolutional neural networks. *arXiv preprint arXiv:1705.10545*, 2017.
- [138] Heung-Il Suk, Chong-Yaw Wee, Seong-Whan Lee, and Dinggang Shen. State-space model with deep learning for functional dynamics estimation in resting-state fmri. *NeuroImage*, 129:292–307, 2016.
- [139] Ian J Goodfellow, Jonathon Shlens, and Christian Szegedy. Explaining and harnessing adversarial examples. *arXiv preprint arXiv:1412.6572*, 2014.
- [140] Jiawei Su, Danilo Vasconcellos Vargas, and Sakurai Kouichi. One pixel attack for fooling deep neural networks. *arXiv preprint arXiv:1710.08864*, 2017.
- [141] Matthew D Zeiler and Rob Fergus. Visualizing and understanding convolutional networks. In *European conference on computer vision*, pages 818–833. Springer, 2014.
- [142] Bratislav Mišić, Richard F Betzel, Marcel A De Reus, Martijn P Van Den Heuvel, Marc G Berman, Anthony R McIntosh, and Olaf Sporns. Network-level structure-function relationships in human neocortex. *Cerebral Cortex*, 26(7):3285–3296, 2016.
- [143] Maarten Mennes, Bharat B Biswal, F Xavier Castellanos, and Michael P Milham. Making data sharing work: The fcp/indi experience. *Neuroimage*, 82:683–691, 2013.
- [144] Chao-Gan Yan, R Cameron Craddock, Xi-Nian Zuo, Yu-Feng Zang, and Michael P Milham. Standardizing the intrinsic brain: towards robust measurement of inter-individual variation in 1000 functional connectomes. *Neuroimage*, 80:246–262, 2013.
- [145] Yves Frégnac and Gilles Laurent. Neuroscience: Where is the brain in the human brain project? *Nature*, 513:27–29, 2014.
- [146] David C Van Essen, Stephen M Smith, Deanna M Barch, Timothy EJ Behrens, Essa Yacoub, Kamil Ugurbil, WU-Minn HCP Consortium, et al. The wu-minn human connectome project: an overview. *Neuroimage*, 80:62–79, 2013.

- [147] Matthew F Glasser, Stamatios N Sotiropoulos, J Anthony Wilson, Timothy S Coalson, Bruce Fischl, Jesper L Andersson, Junqian Xu, Saad Jbabdi, Matthew Webster, Jonathan R Polimeni, et al. The minimal preprocessing pipelines for the human connectome project. *Neuroimage*, 80:105–124, 2013.
- [148] Yashar Behzadi, Khaled Restom, Joy Liau, and Thomas T Liu. A component based noise correction method (compcor) for bold and perfusion based fmri. *Neuroimage*, 37(1):90–101, 2007.
- [149] Klaus H Fritzsche, Peter F Neher, Ignaz Reicht, Thomas van Bruggen, Caspar Goch, Marco Reisert, Marco Nolden, Sascha Zelzer, Hans-Peter Meinzer, Bram Stieltjes, et al. Mitk diffusion imaging. *Methods of information in medicine*, 51(5):441, 2012.
- [150] Bernard Ng, Gaël Varoquaux, Jean Baptiste Poline, and Bertrand Thirion. Implications of inconsistencies between fmri and dmri on multimodal connectivity estimation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 652–659. Springer, 2013.
- [151] Xiaoying Tang, Shoko Yoshida, John Hsu, Thierry AGM Huisman, Andreia V Faria, Kenichi Oishi, Kwame Kutten, Andrea Poretti, Yue Li, Michael I Miller, et al. Multi-contrast multi-atlas parcellation of diffusion tensor imaging of the human brain. *PloS one*, 9(5):e96985, 2014.
- [152] Xiangyu Long, Dominique Goltz, Daniel S Margulies, Till Nierhaus, and Arno Villringer. Functional connectivity-based parcellation of the human sensorimotor cortex. *European Journal of Neuroscience*, 39(8):1332–1342, 2014.
- [153] David Moreno-Dominguez, Alfred Anwander, and Thomas R Knösche. A hierarchical method for whole-brain connectivity-based parcellation. *Human brain mapping*, 35(10):5000–5025, 2014.
- [154] Thomas Blumensath, Saad Jbabdi, Matthew F Glasser, David C Van Essen, Kamil Ugurbil, Timothy EJ Behrens, and Stephen M Smith. Spatially constrained hierarchical parcellation of the brain with resting-state fmri. *Neuroimage*, 76:313–324, 2013.
- [155] Xilin Shen, Fuyuze Tokoglu, Xenios Papademetris, and R Todd Constable. Groupwise whole-brain parcellation from resting-state fmri data for network node identification. *Neuroimage*, 82:403–415, 2013.

- [156] Gagan S Wig, Timothy O Laumann, Alexander L Cohen, Jonathan D Power, Steven M Nelson, Matthew F Glasser, Francis M Miezin, Abraham Z Snyder, Bradley L Schlaggar, and Steven E Petersen. Parcellating an individual subject’s cortical and subcortical brain structures using snowball sampling of resting-state correlations. *Cerebral cortex*, page bht056, 2013.
- [157] Ulrike Von Luxburg. A tutorial on spectral clustering. *Statistics and computing*, 17(4):395–416, 2007.
- [158] Lihi Zelnik-Manor and Pietro Perona. Self-tuning spectral clustering. In *Advances in neural information processing systems*, pages 1601–1608, 2004.
- [159] Lee R Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.
- [160] Harold W Kuhn. The hungarian method for the assignment problem. *Naval Research Logistics (NRL)*, 2(1-2):83–97, 1955.
- [161] Xilin Shen, Xenophon Papademetris, and R Todd Constable. Graph-theory based parcellation of functional subunits in the brain from resting-state fmri data. *Neuroimage*, 50(3):1027–1035, 2010.
- [162] James Munkres. Algorithms for the assignment and transportation problems. *Journal of the Society for Industrial & Applied Mathematics*, 5(1):32–38, 1957.
- [163] Evan M Gordon, Timothy O Laumann, Babatunde Adeyemo, Jeremy F Huckins, William M Kelley, and Steven E Petersen. Generation and evaluation of a cortical area parcellation from resting-state correlations. *Cerebral cortex*, 26(1):288–303, 2014.
- [164] R Cameron Craddock, G Andrew James, Paul E Holtzheimer, Xiaoping P Hu, and Helen S Mayberg. A whole brain fmri atlas generated via spatially constrained spectral clustering. *Human brain mapping*, 33(8):1914–1928, 2012.
- [165] Jiaojian Wang, Yong Yang, Lingzhong Fan, Jinping Xu, Changhai Li, Yong Liu, Peter T Fox, Simon B Eickhoff, Chunshui Yu, and Tianzi Jiang. Convergent functional architecture of the superior parietal lobule unraveled with multimodal neuroimaging approaches. *Human brain mapping*, 36(1):238–257, 2015.

- [166] Jason W Bohland, Hemant Bokil, Cara B Allen, and Partha P Mitra. The brain atlas concordance problem: quantitative comparison of anatomical parcellations. *PloS one*, 4(9):e7200, 2009.
- [167] Qianjun Xu, Marie Desjardins, and Kiri Wagstaff. Constrained spectral clustering under a local proximity structure assumption. In *In FLAIRS Conference*. Citeseer, 2005.
- [168] Bertrand Thirion, Guillaume Flandin, Philippe Pinel, Alexis Roche, Philippe Ciuciu, and Jean-Baptiste Poline. Dealing with the shortcomings of spatial normalization: Multi-subject parcellation of fmri datasets. *Human brain mapping*, 27(8):678–693, 2006.
- [169] Bertrand Thirion, Gaël Varoquaux, Elvis Dohmatob, and Jean-Baptiste Poline. Which fmri clustering gives good brain parcellations? *Frontiers in neuroscience*, 8, 2014.
- [170] Simon B Eickhoff, Klaas E Stephan, Hartmut Mohlberg, Christian Grefkes, Gereon R Fink, Katrin Amunts, and Karl Zilles. A new spm toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage*, 25(4):1325–1335, 2005.
- [171] Martijn Van Den Heuvel, Rene Mandl, and Hilleke Hulshoff Pol. Normalized cut group clustering of resting-state fmri data. *PloS one*, 3(4):e2001, 2008.
- [172] Nathalie Tzourio-Mazoyer, Brigitte Landeau, Dimitri Papathanassiou, Fabrice Crivello, Olivier Etard, Nicolas Delcroix, Bernard Mazoyer, and Marc Joliot. Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *Neuroimage*, 15(1):273–289, 2002.
- [173] John Wiley and Sons. Motor map, 2011.
- [174] Sebastian Bludau, Simon B Eickhoff, Hartmut Mohlberg, Svenja Caspers, Angela R Laird, Peter T Fox, Axel Schleicher, Karl Zilles, and Katrin Amunts. Cytoarchitecture, probability maps and functions of the human frontal pole. *Neuroimage*, 93:260–275, 2014.
- [175] Maxime Descoteaux, Rachid Deriche, Thomas R Knosche, and Alfred Anwander. Deterministic and probabilistic tractography based on complex fibre orientation distributions. *IEEE transactions on medical imaging*, 28(2):269–286, 2009.

- [176] Emmanuel Caruyer, Christophe Lenglet, Guillermo Sapiro, and Rachid Deriche. Design of multishell sampling schemes with uniform coverage in diffusion mri. *Magnetic resonance in medicine*, 69(6):1534–1540, 2013.
- [177] Lipeng Ning, Kawin Setsompop, Oleg Michailovich, Nikos Makris, Martha E Shenton, Carl-Fredrik Westin, and Yogesh Rathi. A joint compressed-sensing and super-resolution approach for very high-resolution diffusion imaging. *NeuroImage*, 125:386–400, 2016.
- [178] Emmanuel J Candes and Benjamin Recht. Exact low-rank matrix completion via convex optimization. In *Communication, Control, and Computing, 2008 46th Annual Allerton Conference on*, pages 806–812. IEEE, 2008.
- [179] Zaiwen Wen, Wotao Yin, and Yin Zhang. Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm. *Mathematical Programming Computation*, 4(4):333–361, 2012.
- [180] Raghunandan H Keshavan, Andrea Montanari, and Sewoong Oh. Matrix completion from a few entries. *IEEE Transactions on Information Theory*, 56(6):2980–2998, 2010.
- [181] Yangyang Xu, Wotao Yin, Zaiwen Wen, and Yin Zhang. An alternating direction algorithm for matrix completion with nonnegative factors. *Frontiers of Mathematics in China*, 7(2):365–384, 2012.
- [182] Yonghui Li, Yong Liu, Jun Li, Wen Qin, Kuncheng Li, Chunshui Yu, and Tianzi Jiang. Brain anatomical network and intelligence. *PLoS Comput Biol*, 5(5):e1000395, 2009.
- [183] Laura Balzano, Robert Nowak, and Benjamin Recht. Online identification and tracking of subspaces from highly incomplete information. In *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*, pages 704–711. IEEE, 2010.
- [184] Rong-En Fan, Kai-Wei Chang, Cho-Jui Hsieh, Xiang-Rui Wang, and Chih-Jen Lin. Liblinear: A library for large linear classification. *Journal of machine learning research*, 9(Aug):1871–1874, 2008.
- [185] Klaus Maier-Hein, Peter Neher, Jean-Christophe Houde, Marc-Alexandre Cote, Eleftherios Garyfallidis, Jidan Zhong, Maxime Chamberland, Fang-Cheng Yeh, Ying Chia Lin, Qing Ji, et al. Tractography-based connectomes are dominated by false-positive connections. *bioRxiv*, page 084137, 2016.

- [186] Hojjatollah Azadbakht, Laura M Parkes, Hamied A Haroon, Mark Augath, Nikos K Logothetis, Alex de Crespigny, Helen E D’Arceuil, and Geoffrey JM Parker. Validation of high-resolution tractography against in vivo tracing in the macaque visual cortex. *Cerebral Cortex*, 25(11):4299–4309, 2015.
- [187] Chad J Donahue, Stamatios N Sotiropoulos, Saad Jbabdi, Moises Hernandez-Fernandez, Timothy E Behrens, Tim B Dyrby, Timothy Coalson, Henry Kennedy, Kenneth Knoblauch, David C Van Essen, et al. Using diffusion tractography to predict cortical connection strength and distance: a quantitative comparison with tracers in the monkey. *Journal of Neuroscience*, 36(25):6758–6770, 2016.
- [188] Sophie Achard, Jean-François Coeurjolly, Romain Marcillaud, and Jonas Richiardi. fmri functional connectivity estimators robust to region size bias. In *Statistical Signal Processing Workshop*, pages 813–816. IEEE, 2011.
- [189] Alex Fornito, Andrew Zalesky, and Edward T Bullmore. Network scaling effects in graph analytic studies of human resting-state fmri data. *Frontiers in systems neuroscience*, 4, 2010.
- [190] Tamás Spisák, András Jakab, Sándor A Kis, Gábor Opposits, Csaba Aranyi, Ervin Berényi, and Miklós Emri. Voxel-wise motion artifacts in population-level whole-brain connectivity analysis of resting-state fmri. *PloS one*, 9(9):e104947, 2014.
- [191] David K Hammond, Pierre Vandergheynst, and Rémi Gribonval. Wavelets on graphs via spectral graph theory. *Applied and Computational Harmonic Analysis*, 30(2):129–150, 2011.
- [192] Ed Bullmore and Olaf Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature reviews. Neuroscience*, 10(3):186, 2009.
- [193] Jianwei Niu, Jinyang Fan, and Ivan Stojmenovic. Jlmc: A clustering method based on jordan-form of laplacian-matrix. In *Performance Computing and Communications Conference*, pages 1–8. IEEE, 2014.
- [194] Santo Fortunato and Marc Barthelemy. Resolution limit in community detection. *Proceedings of the National Academy of Sciences*, 104(1):36–41, 2007.

- [195] David Meunier, Renaud Lambiotte, Alex Fornito, Karen D Ersche, and Edward T Bullmore. Hierarchical modularity in human brain functional networks. *Frontiers in neuroinformatics*, 3, 2009.
- [196] David Meunier, Renaud Lambiotte, and Edward T Bullmore. Modular and hierarchically modular organization of brain networks. *Frontiers in neuroscience*, 4, 2010.
- [197] Roger Guimera and Luis A Nunes Amaral. Functional cartography of complex metabolic networks. *nature*, 433(7028):895, 2005.
- [198] Carlo Nicolini, Cécile Bordier, and Angelo Bifone. Community detection in weighted brain connectivity networks beyond the resolution limit. *Neuroimage*, 146:28–39, 2017.
- [199] Leo Grady. Random walks for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 28(11):1768–1783, 2006.
- [200] Leo Grady. Multilabel random walker image segmentation using prior models. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 763–770. IEEE, 2005.
- [201] Carlo Nicolini and Angelo Bifone. Modular structure of brain functional networks: breaking the resolution limit by surprise. *Scientific reports*, 6, 2016.
- [202] Ian H Witten, Eibe Frank, Mark A Hall, and Christopher J Pal. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2016.
- [203] Burak Yoldemir, Bernard Ng, and Rafeef Abugharbieh. Coupled stable overlapping replicator dynamics for multimodal brain subnetwork identification. In *International Conference on Information Processing in Medical Imaging*, pages 770–781. Springer, 2015.
- [204] Jonas Richiardi, Andre Altmann, Anna-Clare Milazzo, Catie Chang, M Mallar Chakravarty, Tobias Banaschewski, Gareth J Barker, Arun LW Bokde, Uli Bromberg, Christian Büchel, et al. Correlated gene expression supports synchronous activity in brain networks. *Science*, 348(6240):1241–1244, 2015.

- [205] Luca Dodero, Alessandro Gozzi, Adam Liska, Vittorio Murino, and Diego Sona. Group-wise functional community detection through joint laplacian diagonalization. In *MICCAI (2)*, pages 708–715, 2014.
- [206] Michael An, Hon Pong Ho, Lawrence Staib, Kevin Pelphrey, and James Duncan. Multimodal mri analysis of brain subnetworks in autism using multi-view em. In *Signals, Systems and Computers (ASILOMAR), 2010 Conference Record of the Forty Fourth Asilomar Conference on*, pages 786–789. IEEE, 2010.
- [207] Bo Wang, Jiayan Jiang, Wei Wang, Zhi-Hua Zhou, and Zhuowen Tu. Unsupervised metric fusion by cross diffusion. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 2997–3004. IEEE, 2012.
- [208] Abhishek Kumar, Piyush Rai, and Hal Daume. Co-regularized multi-view spectral clustering. In *Advances in neural information processing systems*, pages 1413–1421, 2011.
- [209] Christopher J Honey, Rolf Kötter, Michael Breakspear, and Olaf Sporns. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Sciences*, 104(24):10240–10245, 2007.
- [210] Michael D Fox, Abraham Z Snyder, Justin L Vincent, and Marcus E Raichle. Intrinsic fluctuations within cortical systems account for intertrial variability in human behavior. *Neuron*, 56(1):171–184, 2007.
- [211] Michael W Cole, Danielle S Bassett, Jonathan D Power, Todd S Braver, and Steven E Petersen. Intrinsic and task-evoked network architectures of the human brain. *Neuron*, 83(1):238–251, 2014.
- [212] Micaela Y Chan, Fahd H Alhazmi, Denise C Park, Neil K Savalia, and Gagan S Wig. Resting-state network topology differentiates task signals across the adult life span. *Journal of Neuroscience*, 37(10):2734–2745, 2017.
- [213] Bumhee Park, Dae-Shik Kim, and Hae-Jeong Park. Graph independent component analysis reveals repertoires of intrinsic network components in the human brain. *PloS one*, 9(1):e82873, 2014.
- [214] Maarten Mennes, Clare Kelly, Xi-Nian Zuo, Adriana Di Martino, Bharat B Biswal, F Xavier Castellanos, and Michael P Milham. Inter-individual

differences in resting-state functional connectivity predict task-induced bold activity. *Neuroimage*, 50(4):1690–1701, 2010.

- [215] Maxwell A Bertolero, BT Thomas Yeo, and Mark D’Esposito. The modular and integrative functional architecture of the human brain. *Proceedings of the National Academy of Sciences*, 112(49):E6798–E6807, 2015.
- [216] Angela R Laird, P Mickle Fox, Simon B Eickhoff, Jessica A Turner, Kimberly L Ray, D Reese McKay, David C Glahn, Christian F Beckmann, Stephen M Smith, and Peter T Fox. Behavioral interpretations of intrinsic connectivity networks. *Journal of cognitive neuroscience*, 23(12):4022–4037, 2011.
- [217] M-Marsel Mesulam. From sensation to cognition. *Brain: a journal of neurology*, 121(6):1013–1052, 1998.
- [218] Manlio De Domenico. Multilayer modeling and analysis of human brain networks. *Giga Science*, 6(5):1–8, 2017.
- [219] Sarah Feldt Muldoon and Danielle S Bassett. Network and multilayer network approaches to understanding human brain dynamics. *Philosophy of Science*, 83(5):710–720, 2016.
- [220] Manlio De Domenico, Shuntaro Sasai, and Alex Arenas. Mapping multiplex hubs in human functional brain networks. *Frontiers in neuroscience*, 10, 2016.
- [221] Denny Zhou, Jiayuan Huang, and Bernhard Schölkopf. Learning with hypergraphs: Clustering, classification, and embedding. In *Advances in neural information processing systems*, pages 1601–1608, 2007.
- [222] Pawe Dotko, Kathryn Hess, Ran Levi, Max Nolte, Michael Reimann, Martina Scolamiero, Katharine Turner, Eilif Muller, and Henry Markram. Topological analysis of the connectome of digital reconstructions of neural microcircuits. *arXiv preprint arXiv:1601.01580*, 2016.
- [223] Chen Zu, Yue Gao, Brent Munsell, Minjeong Kim, Ziwen Peng, Yingying Zhu, Wei Gao, Daoqiang Zhang, Dinggang Shen, and Guorong Wu. Identifying high order brain connectome biomarkers via learning on hypergraph. In *International Workshop on Machine Learning in Medical Imaging*, pages 1–9. Springer, 2016.

- [224] Brent C Munsell, Guorong Wu, Yue Gao, Nicholas Desisto, and Martin Styner. Identifying relationships in functional and structural connectome data using a hypergraph learning method. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 9–17. Springer, 2016.
- [225] Mechthild Stoer and Frank Wagner. A simple min-cut algorithm. *Journal of the ACM (JACM)*, 44(4):585–591, 1997.
- [226] Lars Hagen and Andrew B Kahng. New spectral methods for ratio cut partitioning and clustering. *IEEE transactions on computer-aided design of integrated circuits and systems*, 11(9):1074–1085, 1992.
- [227] Jianbo Shi and Jitendra Malik. Normalized cuts and image segmentation. *IEEE Transactions on pattern analysis and machine intelligence*, 22(8):888–905, 2000.
- [228] Karl J Friston, Andrew P Holmes, Keith J Worsley, J-P Poline, Chris D Frith, and Richard SJ Frackowiak. Statistical parametric maps in functional imaging: a general linear approach. *Human brain mapping*, 2(4):189–210, 1994.
- [229] Bernard Ng, Ghassan Hamarneh, and Rafeef Abugharbieh. Modeling brain activation in fmri using group mrf. *IEEE transactions on medical imaging*, 31(5):1113–1123, 2012.
- [230] Thomas Nichols and Satoru Hayasaka. Controlling the familywise error rate in functional neuroimaging: a comparative review. *Statistical methods in medical research*, 12(5):419–446, 2003.
- [231] Peter J Mucha, Thomas Richardson, Kevin Macon, Mason A Porter, and Jukka-Pekka Onnela. Community structure in time-dependent, multiscale, and multiplex networks. *science*, 328(5980):876–878, 2010.
- [232] Jarmo Heinonen, Jussi Numminen, Yevhen Hlushchuk, Henrik Antell, Vesa Taatila, and Jyrki Suomala. Default mode and executive networks areas: Association with the serial order in divergent thinking. *PloS one*, 11(9):e0162234, 2016.
- [233] Caroline J Charpentier, Benedetto De Martino, Alena L Sim, Tali Sharot, and Jonathan P Roiser. Emotion-induced loss aversion and striatal-amygdala coupling in low-anxious individuals. *Social cognitive and affective neuroscience*, 11(4):569–579, 2015.

- [234] Saskia Koehler, Smadar Ovadia-Caro, Elke van der Meer, Arno Villringer, Andreas Heinz, Nina Romanczuk-Seiferth, and Daniel S Margulies. Increased functional connectivity between prefrontal cortex and reward system in pathological gambling. *PLoS One*, 8(12):e84565, 2013.
- [235] Eve Helen Limbrick-Oldfield, I Mick, RE Cocks, John McGonigle, SP Sharman, Anthony P Goldstone, PRA Stokes, Adam Waldman, D Erritzoe, H Bowden-Jones, et al. Neural substrates of cue reactivity and craving in gambling disorder. *Translational psychiatry*, 7(1):e992, 2017.
- [236] Josiah K Leong, Franco Pestilli, Charlene C Wu, Gregory R Samanez-Larkin, and Brian Knutson. White-matter tract connecting anterior insula to nucleus accumbens correlates with reduced preference for positively skewed gambles. *Neuron*, 89(1):63–69, 2016.
- [237] Franco Cauda, Federico D’agata, Katiuscia Sacco, Sergio Duca, Giuliano Geminiani, and Alessandro Vercelli. Functional connectivity of the insula in the resting brain. *Neuroimage*, 55(1):8–23, 2011.
- [238] Anand Venkatraman, Brian L Edlow, and Mary Helen Immordino-Yang. The brainstem in emotion: A review. *Frontiers in neuroanatomy*, 11, 2017.
- [239] Yu Ohmura, Takeshi Izumi, Taku Yamaguchi, Iku Tsutsui-Kimura, Takayuki Yoshida, and Mitsuhiro Yoshioka. The serotonergic projection from the median raphe nucleus to the ventral hippocampus is involved in the retrieval of fear memory through the corticotropin-releasing factor type 2 receptor. *Neuropsychopharmacology*, 35(6):1271, 2010.
- [240] David H Zald. The human amygdala and the emotional evaluation of sensory stimuli. *Brain Research Reviews*, 41(1):88–123, 2003.
- [241] Adrian W Laxton, Joseph S Neimat, Karen D Davis, Thilo Womelsdorf, William D Hutchison, Jonathan O Dostrovsky, Clement Hamani, Helen S Mayberg, and Andres M Lozano. Neuronal coding of implicit emotion categories in the subcallosal cortex in patients with depression. *Biological psychiatry*, 74(10):714–719, 2013.
- [242] Sara M Levens, Orrin Devinsky, and Elizabeth A Phelps. Role of the left amygdala and right orbital frontal cortex in emotional interference resolution facilitation in working memory. *Neuropsychologia*, 49(12):3201–3212, 2011.

- [243] Tömme Noesselt, Nadim Jon Shah, and Lutz Jäncke. Top-down and bottom-up modulation of language related areas—an fmri study. *BMC neuroscience*, 4(1):13, 2003.
- [244] Martin Meyer, Kai Alter, and Angela Friederici. Functional mr imaging exposes differential brain responses to syntax and prosody during auditory sentence comprehension. *Journal of Neurolinguistics*, 16(4):277–300, 2003.
- [245] Jierui Xie, Stephen Kelley, and Boleslaw K Szymanski. Overlapping community detection in networks: The state-of-the-art and comparative study. *Acm computing surveys (csur)*, 45(4):43, 2013.
- [246] Gergely Palla, Imre Derényi, Illés Farkas, and Tamás Vicsek. Uncovering the overlapping community structure of complex networks in nature and society. *arXiv preprint physics/0506133*, 2005.
- [247] TS Evans and Renaud Lambiotte. Line graphs, link partitions, and overlapping communities. *Physical Review E*, 80(1):016105, 2009.
- [248] Yong-Yeol Ahn, James P Bagrow, and Sune Lehmann. Link communities reveal multiscale complexity in networks. *arXiv preprint arXiv:0903.3178*, 2009.
- [249] Shihua Zhang, Rui-Sheng Wang, and Xiang-Sun Zhang. Identification of overlapping community structure in complex networks using fuzzy c-means clustering. *Physica A: Statistical Mechanics and its Applications*, 374(1):483–490, 2007.
- [250] Fan Ding, Zhigang Luo, Jinlong Shi, and Xiaoyong Fang. Overlapping community detection by kernel-based fuzzy affinity propagation. In *Intelligent Systems and Applications (ISA), 2010 2nd International Workshop on*, pages 1–4. IEEE, 2010.
- [251] Andrea Lancichinetti, Santo Fortunato, and János Kertész. Detecting the overlapping and hierarchical community structure in complex networks. *New Journal of Physics*, 11(3):033015, 2009.
- [252] Xiaodan Yan, Stephen Kelley, Mark Goldberg, and Bharat B Biswal. Detecting overlapped functional clusters in resting state fmri with connected iterative scan: a graph theory based clustering algorithm. *Journal of neuroscience methods*, 199(1):108–118, 2011.

- [253] Burak Yoldemir, Bernard Ng, and Rafeef Abugharbieh. Stable overlapping replicator dynamics for brain community detection. *IEEE transactions on medical imaging*, 35(2):529–538, 2016.
- [254] Andrea Lancichinetti, Filippo Radicchi, José J Ramasco, and Santo Fortunato. Finding statistically significant communities in networks. *PloS one*, 6(4):e18961, 2011.
- [255] Andrea Torsello, Samuel Rota Buló, and Marcello Pelillo. Beyond partitions: Allowing overlapping groups in pairwise clustering. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–4. IEEE, 2008.
- [256] Randy L Buckner, Jorge Sepulcre, Tanveer Talukdar, Fenna M Krienen, Hesheng Liu, Trey Hedden, Jessica R Andrews-Hanna, Reisa A Sperling, and Keith A Johnson. Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to alzheimer’s disease. *Journal of Neuroscience*, 29(6):1860–1873, 2009.
- [257] R GeethaRamani and K Sivaselvi. Human brain hubs (provincial and connector) identification using centrality measures. In *Recent Trends in Information Technology (ICRTIT), 2014 International Conference on*, pages 1–6. IEEE, 2014.
- [258] Yong He, Zhang Chen, and Alan Evans. Structural insights into aberrant topological patterns of large-scale cortical networks in alzheimer’s disease. *Journal of Neuroscience*, 28(18):4756–4766, 2008.
- [259] Quintino R Mano, Colin Humphries, Rutvik H Desai, Mark S Seidenberg, David C Osmon, Ben C Stengel, and Jeffrey R Binder. The role of left occipitotemporal cortex in reading: reconciling stimulus, task, and lexicality effects. *Cerebral Cortex*, 23(4):988–1001, 2012.
- [260] Je-Yeon Yun, Jae-Chang Kim, Jeonghun Ku, Jung-Eun Shin, Jae-Jin Kim, and Soo-Hee Choi. The left middle temporal gyrus in the middle of an impaired social-affective communication network in social anxiety disorder. *Journal of Affective Disorders*, 214:53–59, 2017.
- [261] Joana Cabral, Morten Kringelbach, and Gustavo Deco. Functional connectivity dynamically evolves on multiple time-scales over a static structural connectome: Models and mechanisms. *NeuroImage*, 2017.

- [262] Markus Oberweger, Paul Wohlhart, and Vincent Lepetit. Training a feedback loop for hand pose estimation. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3316–3324, 2015.
- [263] Sara Sabour, Nicholas Frosst, and Geoffrey E Hinton. Dynamic routing between capsules. In *Advances in Neural Information Processing Systems*, pages 3857–3867, 2017.
- [264] Carole H Sudre, Wenqi Li, Tom Vercauteren, Sebastien Ourselin, and M Jorge Cardoso. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pages 240–248. Springer, 2017.
- [265] Michael M Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. Geometric deep learning: going beyond euclidean data. *arXiv preprint arXiv:1611.08097*, 2016.