

A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:

<http://wrap.warwick.ac.uk/88528>

Copyright and reuse:

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it.

Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk



**Application of Graph Theoretical Methods to The
Functional Connectome of Human Brain**

by

Soroosh Afyouni

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Institute of Digital Healthcare

September 2016

THE UNIVERSITY OF
WARWICK

Contents

List of Tables	v
List of Figures	vi
Acknowledgments	xii
Declarations	xv
Abstract	xvi
Chapter 1 Introduction	1
1.1 Thesis Objectives and Organisation	2
1.2 Main Contributions	4
Chapter 2 Background:	
From Cajal to Connectome	6
2.1 Human Brain Neuroimaging	6
2.2 Complex Networks	9
2.3 Network Neuroscience	15
2.3.1 Microscale Connectome: Axon Terminals and Dendrites	16
2.3.2 Mesoscale Connectome: Neurons and Minicolumns	16
2.3.3 Macroscale Connectome: Regions of Interest and Pathways	17
2.3.4 Reproducibility and the dead Salmon	21
2.3.5 Summary	22
Chapter 3 Topological Changes in a Network of the Functional Connectivity of the Human Brain in Schizophrenia	23
3.1 Methods	24
3.1.1 Constructing Network of Functional Connectivity	24
3.1.2 Network-based statistics	25

3.1.3	Integration Measures	26
3.1.4	Segregation Measures	26
3.1.5	Network null models	29
3.1.6	Small-worldness	30
3.1.7	Cost-Efficiency	30
3.1.8	Centrality Measures	31
3.1.9	Rich Club	32
3.1.10	Disruption Plots	34
3.2	Data	34
3.2.1	Sample	34
3.2.2	Resting-State fMRI Acquisition and Preprocessing	35
3.3	Results	35
3.3.1	Changes in Edge Weights During Schizophrenia	35
3.3.2	Cost-Efficiency	37
3.3.3	Global Measures: Integration, Segregation and Complexity	38
3.3.4	Centrality Measures in Healthy and Schizophrenia	41
3.3.5	Disruption plots of local centrality measures	44
3.3.6	Changes in Modular Structure during Schizophrenia	45
3.3.7	Changes in Rich Club Organization in Schizophrenia	49
3.4	Discussion	55
3.4.1	Summary	58

Chapter 4 Understanding a Rich Club with a Stochastic Block Model:

	A Connectome Study of Resting-State Schizophrenia	60
4.1	Methods	62
4.1.1	Stochastic Block Model (SBM)	62
4.1.2	Generalised Linear Stochastic Block Model (GL-SBM)	65
4.1.3	Rich Block	67
4.1.4	Degree Analysis	74
4.2	Simulation Methods	75
4.2.1	Validation Methods	75
4.3	Data	77
4.4	Results	77
4.4.1	Simulations	77
4.4.2	Het-SBM on rs-fMRI of Healthy and Schizophrenic Subjects	80
4.4.3	A Rich Block for rs-fMRI Schizophrenia	82
4.4.4	Degree Analysis on rsfMRI Schizophrenia	86

4.5	Discussion	89
4.6	Summary	92

Chapter 5 Impact of Autocorrelation on Topological Features of Functional Connectivity of the Human Brain 93

5.1	Methods	96
5.1.1	Functional Connectivity as a Correlation Matrix	96
5.1.2	Correction Factor	97
5.1.3	Naive Correction	97
5.1.4	Autoregression and Autocorrelation	98
5.1.5	Bartlett's Correction - Markov Process	98
5.1.6	Bartlett's Correction - General Process	99
5.1.7	Family of Heterogeneous Bivariate Corrections (HetBiv)	100
5.1.8	Monte-Carlo Correction	103
5.1.9	Method Comparison	103
5.1.10	Network Thresholding	104
5.1.11	Graph Theoretical Measures	105
5.2	Data	105
5.2.1	Human Connectome Project	105
5.2.2	Autism Brain Imaging Data Exchange	106
5.2.3	Brain Parcellation Schemes	107
5.3	Results	110
5.3.1	Effect of Autocorrelation on Brain Cortical and Sub-cortical Regions	110
5.3.2	Results of the Correction Methods Validation	112
5.3.3	Effect of Autocorrelation Correction on Statistical-Thresholding	114
5.3.4	Effect of Autocorrelation Correction on Graph Theoretical Measures of Statistical-Thresholding	115
5.3.5	Effect of Autocorrelation Correction on Graph Theoretical Measures of Density-Thresholding	122
5.4	Discussion	133
5.5	Summary	137

Chapter 6 Conclusions and Future Work 139

Appendix A Anatomy of the Human Brain 142

Appendix B Rich Block 144

B.1	Derivation of the Variational Bound	144
-----	---	-----

B.2	Stochastic Blockmodel	145
B.3	Coherent and Combinatory Blocks	147
B.4	Projection of Estimated Het-SBM Block Assignments on the Brain Surface	148
B.5	Rich Block of C.Elegans Micro-sclae Connectome	149
B.6	Degree Exceptionality on Group-Level Het-SBM Parameters	150
Appendix C Heterogeneous Bivariate Estimation of Sample Correlation Variance		152
C.1	Mean and Variance of Inner Product of Two Random Variables	152
C.2	Unibased-HetBiv vs Bartlett's HetBiv	153
C.3	Effect of Curbing in HetBiv Corrections	156
C.4	Effect of Shrinking in HetBiv Corrections	158
Appendix D List of Publications		161
D.1	Chapter 3	161
D.2	Chapter 4	161
D.3	Chapter 5	162
Appendix E Codes		163

List of Tables

3.1	Porportion of nodes in each module that contributes to the RC organisation of healthy subjects	50
3.2	Porportion of nodes in each module that contributes to the RC organisation of Schizophrenic subjects	50
5.1	Disruption indices, as discussed in section 3.1.10, of RSNs of three atlases (Power2011, AAL2 and Yeo17) for the NoGSR case of local measures in Statistical Thresholding.	118
5.2	Disruption indices, as discussed in section 3.1.10, of RSNs of three atlases (Power2011, AAL2 and Yeo17) for the NoGSR case of local measures in Statistical Thresholding.	119
5.3	Disruption indices, as discussed in section 3.1.10, of RSNs of three atlases (Power2011, AAL2 and Yeo17) for the GSR case of local measures in Density Thresholding.	126
5.4	Disruption indices, as discussed in section 3.1.10, of RSNs of three atlases (Power2011, AAL2 and Yeo17) for the NoGSR case of local measures in Density Thresholding.	127
B.1	List of blocks and the percentage that each anatomical region participated in them. Coherent blocks are marked by a black asterisk.	147

List of Figures

2.1	A. Headline of the NYT report regarding the Moreno’s presentation (Courtesy of The New York Times). B. Network of blog-to-blog interaction in 2004 US election. Red nodes represent pro-democrat blogs and blue nodes represent pro-republican blogs. Reproduced from Adamic and Glance (2005)	11
2.2	A. a lattice network with Dirac Delta degree distribution B. a smallworld network (SW model). C. a random network with Poisson degree distribution (ER model) D. a Scale-Free network with power-law degree distribution (BA model). Reproduced from Watts and Strogatz (1998) and Barabási <i>et al.</i> (2009).	14
2.3	Cartoon presentation of a small-world network, with modular structure and a strongly connected core which mediates the modules. Reproduced from Bullmore and Sporns (2012a)	15
2.4	Google Trends representing the interest in the topic ‘connectome’ over a period of 12 years (Jan 2004 - Aug 2016). Spikes (shocks) in the search trend are marked with a brief description.	16
2.5	A. Common procedure for constructing the functional connectome. B. Topological description of the human brain, suggesting the human connectome is located between three topological extremes (random, lattice and scale-free). Reproduced from Stam (2014)	18
2.6	Hagmann et al 2008 description of structural connectome: A. Adjacency matrix of the structural connectome of healthy brains. B. Topological presentation of nodes configured based on Kamada-Kawai force-spring layout C. Spatial reconfiguration of figure B with brain region coordinates. Reproduced from Hagmann <i>et al.</i> (2008)	20
3.1	Maslov & Sneppen edge swap procedure on four nodes.	29

3.2	Projections of edges that were found to be significantly decreased during schizophrenia by NBS. Results the NBS over different thresholds were shown in each panel.	36
3.3	Cost efficiency of brain networks in Healthy (Left) and Schizophrenic (Right). Global efficiency of each subject is shown in blue and cost-efficiency is shown in red. The mean cost-efficient density for each group is shown with a vertical black line.	37
3.4	Global measures of segregation, integration and complexity. Panel A. shows the global measures for the level of integration: path length (Left) and normalised path length (Right). Panel B. shows global measures for segregation: clustering coefficient (Left) and normalised clustering coefficient (Right) and maximised modularity index (Bottom) Panel C. shows the small-world index as a measure of complexity	40
3.5	Degree (top) and betweenness (bottom) centrality for averaged across healthy subjects.	42
3.6	Degree (top) and betweenness (bottom) centrality for averaged across Schiz subjects.	43
3.7	Disruption plots of Degree centrality (Left), Betweenness centrality (Middle) and Local Efficiency (Right)	44
3.8	A. Modularity index across different scales for both empirical and randomised networks in healthy (Left). Module configurations on re-arranged group-averaged network of healthy (Right). B. Modularity index across different scales for both empirical and randomised networks in schizophrenia (Left). Module configurations on re-arranged group-averaged network of schizophrenia (Right).	46
3.9	Projection of individual nodes and their module assignments on a transparent surface of the human brain.	47
3.10	A river-plot shows the overlap between modular structure of healthy and schizophrenia subjects.	48
3.11	Similarity between node assignments of modularity algorithm across subjects of healthy (Left) and schizophrenia (Right)	49
3.12	Projection of Rich Club organisation in healthy (green), schizophrenia (blue) and overlaps between two groups (cyan)	50

3.13	Interaction between RC organisation and rest of the network via feeder and local connections. A. Analysis of density between feeder, local and RC nodes as well as between healthy and schizophrenia subjects. B. Analysis of Local efficiency of local, feeder and RC nodes as well as between two groups. C. Analysis of Betweenness Centrality of feeder, local and RC nodes. Red asterisks indicate significant differences.	52
3.14	NRCC of healthy (Top) and schizophrenia (Bottom), as described in section 3.1.9. Nodes that were identified as significantly different were marked with a red asterisk	53
3.15	NRCI of healthy (Top) and schizophrenia (Bottom), as described in section 3.1.9. Nodes which were identified as significantly different were marked with a red asterisk	54
3.16	Similarity between Rich Club Identity matrices across individuals across healthy (Left) and schizophrenia (Right)	55
4.1	Rich Block detection procedure on a toy connectivity rate matrix, $\hat{\Pi}$, and its null model, Π_{Null} , were shown, representing a network of 300 nodes and $Q = 15$. Binomial Mixture densities used to generate degree distributions of two arbitrary RB organisation of 4 (A) and 11 (B) illustrated. The results of RB detection suggest that $H_0 : f^k(\rho_i) > f_{Null}^k(\rho_i)$ is significant, while similar hypothesis fails to be rejected on the 11th RB organisation.	71
4.2	A. Connectivity probability matrix of a toy network with 15 (equal-sized) blocks B. Simulated network. C. shows the RB coefficients, for this network. It is notable how the RB coefficients could differentiate between the blocks with the same expected degree but different patterns of connectivity.	72
4.3	Flowchart of calculation Rich Block and group inference between two groups. Rectangles indicate process, parallelogram shows data. Blue indicate processes that should be under taken on control and red for patient. Also note that colourful sections should be repeated for each individual, however, we only depicted this process for one Π per group.	73
4.4	Shows probability density functions of how the proposed Tensor Null Model behaves in Rich Block organisation.	78

4.5	Shows results of RB estimation for both coefficient and organisation. Each column represent networks with different size and each row represent networks with different block numbers. In each subplot, the left-hand bar represent the coefficient of deterministic (r^2) between the RC and RB coefficients and right-hand bar represent the Adjusted Rand Index between the RC and RB organisations.	79
4.6	Shows results of the Degree Exceptionality on simulated networks. Each column represents networks of different sizes and each row represents networks of different block numbers. The number of hubs in each network was defined as 10% of the number of blocks.	80
4.7	Connection probability matrix of control A and Schizophrenic patients B . C and D are sorted connection probability matrices based on the decreasing expected degree of each block.	82
4.8	illustrates a circular layout of the group-averaged probability connection matrix, Π , of the controls A and patients B . Detected RBs are located in the middle of the layout. For the sake of visualisation, we only drew the top 30% of the strongest connections.	83
4.9	Shows the Rich Block Identity (RBI) of controls (top) and patients (bottom). The yellow cells indicate that block q for subject s is a significant Rich Block. Red asterisk indicates statistical differences between two RBIs.	84
4.10	A and C shows Block-wise Rich Block Coefficient (BRBC) in control and patient respectively. B and D shows mean of each column in A and B . Red asterisk shows statistical significance between mean of RB values of two groups.	85
4.11	A . illustrates the organisation of the local and feeder connections along with RB in a synthetic adjacency matrix. B . Illustrates the difference in density of the RB connection individuals in healthy and schizophrenia. C . Shows the feeder density of individuals in healthy and schizophrenia subject groups. D . Shows the local density between individuals of healthy and schizophrenia. The red asterisk indicates the statistical difference.	86
4.12	A and B are binary matrices of the DE effect in healthy and schizophrenia groups. B and D are the expected degree of each block for the individuals in the control and patient group. Red asterisk indicates the statistical difference and red rectangles indicate the degree exceptional blocks.	87
4.13	Illustrates a Bland-Altman plot to investigate hub distribution between healthy and schizophrenic patients. The black dotted line shows the mean of the differences. The grey dotted lines indicate the 1 SD distance from the mean.	88

5.1	A,B,D Group-averaged correlation length for Yeo17, AAL2 and Power2011, respectively. C. Volume of the ROIs across the parcellation schemes.	110
5.2	Association between correlation length and four different parcellation schemes (e.g. Yeo17, AAL2, AAL, CC200) of three different scanning sites (e.g. Yale, NYU, USM and HCP) with and without GSR.	111 111
5.3	KS statistics of a family of HetBiv corrections as well as existing methods for each acquisition site and parcellation scheme. The grey lines represent the reference line.	112
5.4	False-positive rate of family of HetBiv corrections as well as existing methods for each acquisition site and parcellation scheme. The acquisition site has been identified by the titles.	114
5.5	Histogram of the Fisher-transformed correlation coefficients after density thresholding with Naive and HetBiv correction for the degree of freedom.	115
5.6	Network density of networks with Naive and HetBiv corrections. Each subplot shows the results for the choice of GSR and parcellation scheme. The red asterisk shows statistically significant results.	116
5.7	Global graph theoretical measures of networks with Naive and HetBiv corrections. Each subplot shows the results for choice of GSR and parcellation scheme. Red asterisk shows statistically significant results.	117
5.8	Disruption plots of the changes in the local measures (Degree Centrality, Betweenness Centrality and Rich Clubs) of the networks with statistical thresholding due to HetBiv correction. All of the x-axes represent the average of the measure across groups and all of the y-axes represents the differences in the values between the two groups.	123 123
5.9	Change rates between global measure of network with Naive correction and HetBiv correction. Each column is a parcellation scheme and each row is global graph theoretical measure. Red circle indicates the statistical significance between the two correction in case of NoGSR. Black cross indicates the statistical significance between the two correction in case of GSR. Vertical lines show cost-efficient densities over different choices of GSR and correction methods. HetBiv-corrected, GSR (dashed blue); HetBiv-corrected, NoGSR (dashed green); Naive-corrected, GSR (solid black); Naive-corrected, NoGSR (solid red). All of the x-axis shows the density, and all the y-axis shows the Change Rate.	125 . 125

5.10	Disruption plots of changes in local measures (Degree Centrality, Betweenness Centrality and Rich Clubs) of network with density thresholding due to HetBiv correction. All of the x-axes represent the average of the measure across groups and all of the y-axes represents the differences in the values between the two groups.	132
A.1	illustrates the anatomy of the human brain including the lobes and the areas within each lobe.	143
B.1	Visualisation of block on MNI surface of the human brain. For the sake of visualisation, we show only 5 or 6 blocks per each brain surface.	148
B.2	Non-normalised RC (solid grey) and RB with optimal Q (marked blue) and fixed Qs (marked red,green,magenta and black)	150
B.3	Illustrates estimated PDF of the group-averaged connection probability matrix for controls (left) and patients (right). The red asterisk at the top of each Poisson component indicates degree exceptionality. Also the expected degree of each component involved in estimation of the degree distribution of each group is listed respective to their colour.	151
C.1	KS Statistics of HetBiv methods. Each subplot represents a cohort: HCP (Top Left), NYU (Top Right), USM (Bottom Left) and Yale (Bottom Right).	154
C.2	False-positive rates of the HetBiv methods. Blue vertical lines represent the 5% α -level and red vertical lines represent 2.5% α -level.	155
C.3	KS statistics of each curbing coefficient across different acquisition sites and parcellation schemes.	156
C.4	False-positive rate analysis of each curbing coefficient across different acquisition sites and parcellation schemes.	157
C.5	KS statistics of each shrinking coefficient across different acquisition sites and parcellation schemes.	158
C.6	False-positive rate of each shrinking coefficient across different acquisition sites and parcellation schemes.	159

Acknowledgments

First and foremost, I wish to thank my supervisors; Professor Thomas Nichols and Professor Theodoros Arvanitis, not only for their valuable time, ideas and funding, but also for giving me the opportunity to feel like a contributing member of a strong team (which was not always the case). You, consciously and subconsciously, taught me how to be persistent, skeptical and creative throughout the scientific process. Your patience during our Thursday meetings shaped me as a researcher, and I genuinely hope that I was worth the investment.

I would also like to thank Dr Andrew Bagshaw. If it was not for your support and patience during the early stages of my PhD, I would have not been able to accomplish what I presented in this thesis. On these lines, I would also like to thank past and present members of the Multimodal Imaging Group at the University of Birmingham; specially Dr Joanne Hale for her suggestions and contributions. I am also thankful to Dr Dragana Pavlović for her constant patience when walking me through various complex statistical models. Grateful acknowledgements to Professor Edward Bullmore, as well as members of Human Brain Mapping Unit at the University of Cambridge (specially Dr Petra Vértes and Dr Emma Towlson), for your stimulating ideas and suggestions.

My time at the University of Warwick was made enjoyable in large due to the energetic, collaborative and jolly spirit of past and present members of the Institute of Digital Healthcare and the Statistical Neuroimaging Group at the Department of Statistics. You were not only colleagues and officemates, but also friends. Thank you for bearing my often-miserable face and my constant complaints during the difficult period of writing this thesis. I would also like to thank Dr Ahmed Fetit and Dr Sarah Keung for helping me edit this thesis.

I am also grateful for financial support from the Institute of Digital Healthcare,

Division of Brain Sciences at Imperial College London and Guarantors of Brain. I would also like to thank Institute of Advanced Study at University of Warwick that made me sure that I will not be unemployed after submitting this thesis.

Also, I would like to thank Ms Mehrnoush Sarafan, whose kindness, understanding and patience tremendously helped me during my PhD. I should also thank my friends, who I grew up with, but then later scattered across different countries; however distant, you never discounted our friendship. You are a living evidence of a possible, significant correlation between millage and friendship. Thanks for your support and for all the banter.

To mum, dad and Shokouh for everything. During the difficult moments of this PhD, and throughout my whole life, I was constantly reminded that there is a safe haven called *Home*, where I am offered unconditional love and a cup of freshly brewed tea. I am genuinely sorry that these words are my only way of thanking you.

To my friends of misery.

Declarations

I, Soroosh Afyouni, hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in these, or any other Universities. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text.

- Parts of Chapter 3 were presented in Organisation for Human Brain Mapping (OHBM) annual meeting in 2014.
- Chapter 4 is based on a pre-printed manuscript, which will be shortly submitted for publication. Parts of this chapter were presented during OHBM annual meetings in 2015 and 2016.
- Chapter 5 is based on a pre-printed manuscript, which will be shortly submitted for publication. This work is presented as an abstract in OHBM annual meeting in 2016.

Soroosh Afyouni

September 2016

Abstract

During the past decade, there has been a great interest in creating mathematical models to describe the properties of connectivity in the human brain. One of the established tools to describe these interactions among regions of the brain is graph theory. However, graph theoretical methods were mainly designed for the analysis of single network which is problematic for neuroscientists wishing to study groups of subjects. Specifically, studies using the Rich Club (RC) graph measure require cumbersome methods to make statistical inferences. In the first part of this work, we propose a framework to analyse the inter-subject variability in Rich Club organisation. The proposed framework is used to identify the changes in RC coefficient and RC organisation in patients with schizophrenia relative to healthy control. We follow this work by proposing a novel method, named Rich Block (RB), which is a combination of the tradition Rich Club and Stochastic Block Models (SBM). We show that using RBs can not only facilitate a inter-subject statistical inference, it can also account for differences in profile of connectivity, and control for subject-level covariates. We validate the Rich Block approach by simulating networks of different size and structure. We find that RB accurately estimates RC coefficients and RC organisations, specifically, in network with large number of nodes and blocks. With real data we use RB to identify changes in coefficient and organisation of highly connected sub-graphs of hub blocks in schizophrenia. In the final portion of this work, we examine the methods used to define each edge in networks formed from resting-state functional magnetic resonance imaging (rs-fMRI). The standard approach in rs-fMRI is to divide the brain into regions, extract time series, and compute the temporal correlation between each region. These correlations are assumed to follow standard results, when in fact serial autocorrelation in the time series can corrupt these results. While some authors have proposed corrections to account for autocorrelation, they are poorly documented and always assume homogeneity of autocorrelation over brain regions. Thus we propose a method to account for bias in inter-region correlation estimates due to autocorrelation. We develop an exact method and an approximate, more computationally efficient method that adjusts for the sampling variability in the correlation coefficient. We use inter-subject scrambled real-data to validate the proposed methods under a null setting, and intact real-data to examine the impact of our method on graph theoretical measures. We find that the standard methods fail to practically correct the sensitivity and specificity level due to over-simplifying the temporal structure of BOLD time series, while even our approximate method is substantially more accurate.

Chapter 1

Introduction

When observing advances in biology, one could argue that modern neuroscience was ignited by a series of investigations conducted by Santiago Ramon y Cajal in the late 19th Century. Upon analysing thousands of microscopic samples gathered from bird cerebellum and adult vertebrate nervous system, Cajal found that neurons tend to interact with each other via a *contiguous* process (Swanson and Lichtman, 2016). At the time, this contradicted the conventional way of treating brain circuitry as a *continuous* collection of neurons.

Interestingly, long before Cajal's findings, there were reports that suggested the presence of interaction between distinct brain cortex areas and functions. Such relationships were confirmed several years later after a revolution in neuroimaging technologies. Functional localisation of brain activities unveiled that hundreds of thousands of neurons in a spatially constrained area are involved in a task (e.g. motor process) and can be seen as the inception of the notion of segregation in the brain (Edelman, 1987). However, further examination of brain functions, such as attentional mechanisms, showed that carrying out a single task involves several spatially distributed brain areas, which facilitate the functional task through a mediation between them (Edelman, 1993). The aforementioned findings suggest that function of the human brain is built upon a simultaneous coexistence of anatomical *segregation* and functional *integration*.

Emergence of new interdisciplinary fields, such as *complex system sciences*, has motivated neuroscientists to establish a computational framework combining Cajal's doctrine and the segregated/integrated nature of the human brain. Proposing 'Neural Complexity' was one of the earliest attempts to formulate the coexisting anatomical dependence (segregation) and functional independence (integration) (Tononi *et al.*, 1994). Neural Complexity uses Information Theoretical techniques to quantify the presence of the *complexity* in the vertebrate brain, by capturing the interplay between segregation and integration of function and structure. It concludes that the human brain is a *complex system* due to pres-

ence of a balance between these two tendencies. Neural Complexity, later, was elaborated to a more sophisticated quantification tool, called Integrated Information Theory, which attempts to explain the mind-body problem and eventually the hard problem of consciousness (Harnad, 1994).

Despite graph theory being a relatively old branch of mathematics, it is not until long ago that it was recognised as a powerful tool for modelling system complexities. It represents agents of a system as *nodes* and interaction between them as *edges*. Despite its apparent simplicity, graph theory showed that it is able to model extremely complex systems such as, human social behaviour (Fararo and Sunshine, 1964; Klov Dahl *et al.*, 1994; Newman, 2001). Mammalian brains are also not an exception as being one of the most complex yet unknown natural systems. Application of graph theoretical methods has demonstrated strong evidence in modeling the segregation, integration and eventually complexity of the brain networks. Presence of segregation has been confirmed by network clustering techniques, which suggest that the brain networks follow a modular and hierarchical topological structure (Meunier *et al.*, 2010). Also, the presence of relatively short fiber tracts, which mediates the brain regions, confirms the integrative nature of the human brain (Sporns *et al.*, 2004). Modular structures, which are connected via short paths, confirm the constant battle between segregation and integration and, analogous to Neural Complexity, suggest the presence of the ‘small-world’ phenomenon within the human brain (Bassett and Bullmore, 2006).

Throughout the last decade, advancements in field of graph theory have gone far beyond deterministic measures of segregation, integration and complexity. Today, examination of the latent structures within a network, by using stochastic graph theoretical methods, can unveil extensive amounts of information about current network states, as well as future (generative) network behaviour (Karrer and Newman, 2011). Therefore, it is important to maintain a sturdy bridge between the fields of graph theory and neuroscience, aiming to maximise discoveries that can be delivered by computational brain modelling.

1.1 Thesis Objectives and Organisation

The primary aim of this thesis is to contribute to bridging the two disciplines by firstly, showing that graph theory is able to deliver valuable information about the structure and function of the human brain and therefore can be used as a plausible bio-marker to detect the alteration in the brains of patients with neuro-psychiatric disorders. Secondly, demonstrating that newly-proposed stochastic block modelling methods can be used to gain a better insight into the topological features of the human brain. Thirdly, translating the well-established deterministic graph theoretical metrics into the stochastic block modelling

domain aiming to empower the application of graph theory for understanding complex phenomena. Fourthly, tackling the current challenges in constructing brain networks by eliminating the nuisance variances by estimating the effect of correlated noises that exist in time series of brain functional activities.

In chapter 2, we begin by giving a brief description of brain imaging methods, specifically, Functional Magnetic Resonance Imaging. This is followed by a short history and summary of network sciences. Eventually, we give a summary about combination of these two disciplines which is known as 'connectome'. Graph theoretical methods that are mentioned in this chapter are cross-referenced to a deeper theoretical discussion throughout the thesis.

In chapter 3, we demonstrate how the fundamental understanding of the human brain can be examined by graph theoretical measures. Additionally, we investigate how more sophisticated metrics can help us to gain a better understanding of human brain function. Ultimately, we illustrate how graph theory can be considered as a bio-marker to detect the alterations in the brain networks of patients with schizophrenia.

In chapter 4, a newly proposed Stochastic Block Model is used to investigate how this graph theoretical method can be used to describe topological properties of brain functional connectivity. Additionally, we introduce a novel method, Rich Block, to detect highly connected network cores. This method is an interpretation of the traditional Rich Club with the use of Stochastic Block Models. We also introduce a framework to conduct multi-subject statistical inference between patient and healthy control groups. Further to Rich Block, we also propose Degree Exceptionality which can be used to detect the hubs as well as their alterations during the disease. These methods are applied to the functional connectivity of Schizophrenic patients to demonstrate how they can be used as bio-markers to track alterations during the disease.

In chapter 5, we investigate how correlated noise can affect the functional connectivity maps and, consequently, topological description of the human brain. We focus on addressing the dependencies between BOLD signal data-points, due to the presence of autocorrelation effects in BOLD signals. We show that the effect of this specific category of noise on the human cortex is rather heterogeneous. To mitigate this effect we propose a novel technique to estimate the nuisance variance of correlation coefficients, called the Heterogeneous Bivariate Correction Factor (HetBiv-correction). We demonstrate how this technique out-performs existing methods in correcting for the effects of autocorrelation. In addition, we show how HetBiv-correction can be helpful in revealing the topological features of functional connectivity maps.

In conclusion, in chapter 6, we summarise the findings of this work and argue that network neuroscience is, not the only, but one of the more credible methods of describing

the function of the human brain. However, as a young interdisciplinary field, it is subjected to several challenges and limitations and employing it requires careful and rigorous analysis.

1.2 Main Contributions

In this work, we make a number of contribution to the field of network neuroscience:

- Incentivised by lack of statistical techniques to conduct group inference between the Rich Club configuration of functional connectome across subjects, we proposed a statistical framework, which accounts for within group variability in the analysis of traditional Rich Club configuration. This novel approach differentiates between Rich Club coefficients and Rich Club organisation to maximise the detection power between two groups. We used a previously published data set with a new brain parcellation scheme to demonstrate how this technique can be used to fill in the gaps of between-group inference in the detection of changes in patients with schizophrenia. The methods proposed help the network neuroscientists to conduct group inference between the RC organisation and coefficients of the healthy and patient groups.
- As the Rich Club techniques are known to be highly sensitive to the degree of nodes, we combined a traditional Rich Club detection technique with SBMs. This combination, named Rich Block, can be seen as a complimentary method to SBMs as long as they assume that the probability that a node makes a connection follows a Bernoulli distribution. In addition to accounting for high degree centrality in the formation of ‘clubs’, Rich Block also differentiates between nodes of a club which has the same degree centrality but different profile of connectivity.
- In addition to accounting for the profile of connectivity, by using a newly proposed Multi-Subject Stochastic Block Models, Rich Block can also be used to account for within group variability in both coefficients and the organisation of Rich Blocks and therefore facilitates the group inference between groups. Using multi-subject SBM also provides a unique opportunity to account for the effect of nuisance covariates in between group comparisons. Accounting for the covariates helps researcher in the community to conduct more reliable hypothesis tests. We apply Rich Blocks to a group of healthy and schizophrenic subjects and show how the changes in their network cores can be detected by Rich Block.
- Rich Block organisations should be tested against an appropriate null model to ensure that they are not merely due to chance. Motivated by the lack of null models in

the context of SBM, we proposed Tensor Null Models (TNM) which represent a structure-free expected connectivity rate between estimated blocks. This null model is independent of empirical network, and therefore, facilitates the statistical inference between the null RB organisation and the estimated RB organisation.

- Topological properties of the functional connectivity can be drastically influenced by inevitable temporal confounds, such as correlated noise, which can easily jeopardise the independency between observations and inflate the sampling variance. Motivated by the lack of methods for mitigating the effect these temporal confounds, we propose a novel method, named HetBiv-correction, to estimate the sampling variance of correlation coefficients. By employing the Kolmogorov-Smirnov statistics and false-positive rates, we validated the proposed method using inter-subject scrambling method. The results of validation suggest that the HetBiv outperforms, in regards to both shape of the distribution and tails, the existing methods by estimating a more accurate sampling variance. For instance, in the HCP cohort, the KS statistic was found to be 0.003 and the false-positive rate of the proposed methods completely overlaps with both 0.01 and 0.05 α -levels.
- In addition to inflated sampling variance, our investigations suggest that the effect of temporal confounds is not spatially homogeneous across the brain, and existing homogeneous approaches may sacrifice the sensitivity and specificity level for a simpler temporal model. On the contrary, HetBiv-correction estimates this variance for every connections and therefore accounts for the effect of spatial heterogeneity of temporal confounds in functional connectivity of the human brain.

Chapter 2

Background: From Cajal to Connectome

In this chapter, we briefly discuss the advancement in human brain imaging techniques and how they can help to gain a better insight about functional and structural configuration of the human brain. We follow with a brief introduction to graph theory and examples of how it can be used as a mathematical model to describe natural phenomena. Eventually, we discuss the emerging field of network neuroscience, which is centered on employing graph theoretical methods and advancing imaging techniques to deliver a comprehensive description of the human brain.

2.1 Human Brain Neuroimaging

Non-invasive recording of brain neuronal activities is, undoubtedly, one of the most remarkable milestones in neuroscience, aiming to understand the function and structure of the human brain mechanisms. It is now a few decades since the Electroencephalogram (EEG), as the first non-invasive method of recording brain electrophysiology and function, helped neuroscientists to monitor human brain activities. Despite the fine-scale temporal resolution that this recording method provides for investigating the superficial areas of the brain such as motor or visual systems, it was shown that the majority of human brain activities happen in deeper parts of the brain, where it is impractical for EEG techniques to reach them or at least deliver reliable information about them. These limitations encouraged researchers to seek more sophisticated methods of brain structural and functional mapping of the human brain, through biomedical imaging.

In the late 1980's and early 1990's a chain of advancements in non-invasive brain imaging techniques sparked a revolution in the field of neuroscience. These imaging tech-

niques ranged from radiation-based imaging techniques, such as Computer Tomography (CT) and Positron Emission Tomography (PET) scanning, to more advanced imaging methods, such as Magnetic Resonance Imaging (MRI), which can provide high spatial resolution images of the function and structure of the brain (Penny *et al.*, 2011).

Functional Magnetic Resonance Imaging (fMRI), which is the main interest of this thesis, has gained enormous attention since its introduction, in the 1990's. This biomedical imaging method works by monitoring the changes in demand for oxygenated blood of the brain regions. Activation of different brain regions is measured by blood-oxygenation-level-dependent (BOLD) contrasts. BOLD fMRI signals are typically acquired using an echo-planar imaging technique (Westbrook and Roth, 2011) which is formed by applying a Radio Frequency (RF) pulse in the presence of a static magnetic field to selectively excite and encode magnetisation vectors (Lauterbur, 1973). After a fraction of a second, the excitation is turned off and the reaction of water molecules is recorded in a k-space during their relaxation phase. Inverse Fourier transformation then translates the acquired information from the frequency domain to the time domain (Westbrook and Roth, 2011). This procedure is repeated over a specific period of time to record the dynamics of the brain activities. This procedure results in a BOLD signal representing functional activation of a small isomorphic cubical (normally $2^3 mm$) volume of the brain called a *voxel*. It is worth noting that BOLD signals are not direct measures of the brain functional activities but, rather, a representation of its neurovascular coupling into hemodynamic changes (Roche and Commins, 2009).

Although until the early 1990s, fMRI was mainly used as a tool to measure task-dependent activations of the brain, an extraordinary finding reshaped the research on the function of the brain. In 1995, Biswal *et al.* (1995) discovered that task-free brain scans also demonstrate a level of activation in certain frequency bands which is called *resting-state* brain (Raichle and Snyder, 2007). They argued that the majority of the brain regions are constantly active in a frequency band of $0.01 - 0.1 Hz$, which later led to the realisation that the brain energy level increase by only $\approx 5\%$ (Raichle and Snyder, 2007) during a task. The introduction of resting-state brain imaging facilitated the discovery of a range of brain functional networks (RSNs) such as the Saliency Network (Fox *et al.*, 2005) and most importantly the Default Mode Network (DMN) (Sheline *et al.*, 2009), which is known to be active when a subject is awake, but not performing any specific task. The groundbreaking realisation about RSNs roots, down to the fact that the anatomical regions, which are involved in RSNs, are spatially spread around the brain whilst consistently acting together. To date, neuroscientists have discovered more than 7 well-established RSNs, as well as 17 sub-networks from a Visual and Somatomotor network to a Frontoparietal and Limbic network (Yeo *et al.*, 2011). It is also worth noting that despite the reproducible and meaningful

evidence of a resting-state brain, the physiological meaning of a resting-state brain is still the subject of an ongoing debate since there are studies, which show that resting-state oscillations may have been caused by cardiac (Birn *et al.*, 2006) and respiratory oscillations (Chang and Glover, 2009).

Acquired BOLD signals, in the time of scanning, are far from being ready to be considered as a reliable description of brain activities and require undertaking a number of pre-processing steps to remove the variances that may have been merely caused by noise and which confound during the scan. The pre-processing steps can be summarised in five stages.

Slice timing correction: fMRI scans are a collection of 3D snap-shots of the brain for each time-point over the scanning session. Each of the 3D images is reconstructed from a number of 2D slices, however, the timing between these slices may differ between subjects and scan sessions. Slice timing correction aims to temporally align these slices between all subjects (Henson *et al.*, 1999).

Motion Correction: Acquiring each 3D image may take up to a couple of seconds; during this time even a smallest head movement can cause propagative disturbance in the analysis of functional MR images (Van Dijk *et al.*, 2012). Regardless of how a subject can manage to stay still inside the scanner, physiological functions, such as cardiac and respiratory processes, can cause movements between slices, which makes it vital to correct the spatial mis-alignment between the slices. This effect is typically eliminated by a method called Rigid Body Realignment (RBR) (Ashburner *et al.*, 2003). Further to the traditional method of correcting the head movement, there are several complimentary methods for motion correction such as motion scrubbing (Power *et al.*, 2014) FMRIB's ICA-based Xnoiseifier (FIX) ¹ (Salimi-Khorshidi *et al.*, 2014) and spike wavelet analysis (Patel *et al.*, 2014).

Registration: After spatial alignment, each subject's scan remains in its own scanner coordinate space, however, it is important to transform all the scans into the same, standard, coordinate space. Therefore, registration is performed to transform all images to a standard space (Penny *et al.*, 2011). Further to global alignment, it also helps to remove the anatomical difference between subjects such as differences in size and surface of the brain. One of the widely-established examples of standard spaces, which we use throughout this thesis, is formed by averaging anatomical scans of 152 subjects. This standard template is called MNI-152 (which stands for Montreal Neurological Institute) (Collins *et al.*, 1995).

Spatial and Temporal Smoothing: These two classes of smoothing are used to improve the signal-to-noise ratio (SNR) and, therefore, the sensitivity in later statistical inference. Spatial smoothing is performed by applying a 3D Gaussian kernel to the images

¹The term FMRIB is short for Centre for Functional Magnetic Resonance Imaging of The Brain

(Penny *et al.*, 2011). However, the choice of bandwidth of this kernel is crucial as the choice of large bandwidths can blur the image whereas small bandwidths may not be effective Wang *et al.* (2005). Further to spatial smoothing, temporal smoothing was also shown to improve the data quality. Although it is recommended to only apply high-pass filtering and avoid low-pass filtering as there is evidence that information relies on the lower frequencies (Van Essen *et al.*, 2012).

Global Signal Regression (GSR): GSR is, undoubtedly, one of the most controversial topics in fMRI pre-processing. In 2009, GSR seemed to be an intuitive way of removing the effect of natural processes such as cardiac and respiratory functions as well as nuisances caused by white matter and CSF in the BOLD signal (Fox *et al.*, 2009). However, not long after the initial proposal, it was shown that regressing the global signal might remove physiologically meaningful information (Murphy *et al.*, 2009). As GSR is currently a subject of ongoing debate, it is normally advised to report the results for both cases with and without GSR. Moreover, recently, several other methods have been proposed to remove nuisance variances without removing the meaningful physiological information (Carbonell *et al.*, 2011; Behzadi *et al.*, 2007).

For extended discussion about effect of GSR; see chapter 5

Although this work is centred on fMRI, MRI capabilities are far beyond merely functional imaging. One of the well-known methods for structural imaging of the human brain is diffusion-weighted MRI (DWI), which produces contrast in MR images based on random motion of the water molecules, or as conventionally named Brownian Motion (Westbrook and Roth, 2011). This technique facilitates the monitoring of the diffusion of water molecules and consequently detecting the anatomical features of the brain tissues. Today, a special form of DWI, Diffusion Tensor Imaging (DTI), is widely used to detect the axonal fibre tracts within the white matter of the human brain. Other techniques for structural imaging such as conventional (T1 and T2 weighted) images of the human brain also revealed useful information about brain lesions, as well as more sophisticated information on grey matter structure such as cortical thickness. For a detailed discussion about DTI techniques see (Hagmann *et al.*, 2006).

2.2 Complex Networks

In its simplest form, a network is a collection of dots and lines joined together to model a phenomenon. In mathematics, a collection of these joined dots and lines is also called *graph* and the area of studying the interaction between these entities is a comprehensive branch of mathematics called *graph theory*. In graph theory terminology, each of the dots is commonly called a node or a vertex, and each of the lines representing the interaction between nodes is called an edge (Newman, 2002).

The term 'graph' was first coined by Sylvester in a paper published in the American Journal of Mathematics in 1878 titled *On an application of the new atomic theory to the graphical representation of the invariants and covariants of binary quantics* (Sylvester, 1878). However, the notion of representing the entities of a system with nodes and edges travels back to a century before Sylvester's paper, in 1736, where Leonhard Euler, a Swiss mathematician, was trying to solve the then famous Königsberg bridge problem. The problem was about how one can cross all seven bridges in Königsberg without crossing a bridge more than once. Euler formulated the problem as a collection of nodes (shores) and bridges (edges) and found that such a path is impossible (Euler, 1741). His finding was formulated later as the first theorem in graph theory, which was named after himself. An Euler path crosses all the nodes without visiting any edges twice.

Since Euler's time, graph theory has been well established in the modelling of empirical data in different branches of science and technology. One of the many example of this exploitation is power grid networks, which are networks of high-voltage transmission lines that provide long-distance transport of electrical power within and between countries (Dobson *et al.*, 2007). In a power grid network, generating stations and electrical dispatches are analogues to nodes, while transmission lines are represented by edges. Although, abstracting the large-scale power grid to a collection of nodes and edges may seem to be oversimplifying a system with such a scale of complexity, this model can carry extraordinarily complicated information such as the flow of electricity in each submission line in addition to the capacity of each generating/dispatch post which varies across regions. One common incident that can be examined by a power grid network is *cascade failure* of transmission lines, which happens when a failing power line en-routes its electricity flow to other power posts and cause them to overflow and, consequently, fail too. This effect can be propagated in the system and cause major failure in power transmission system. Beside monitoring the network resilience in power grid networks, it also aid the authorities to have a better insight of the network *topology* which is the configuration of nodes and edges regardless of geography.

A more sophisticated example of applying graph theory to empirical networks is social networks, where graph theory attempts to deliver a comprehensible representation of highly complicated human behaviours. Among different disciplines, social sciences, undoubtedly, had the largest contribution to the field of network sciences. This contribution ignited in 1933 with Jacob Moreno, a Romanian social scientist, presented a network of relationships between individuals attending a conference taking place in New York (Moreno, 1934). His work was featured in the New York Times (see Figure 2.1) shortly after his presentation and drew the attention of other researchers to regard graph theory as a viable tool to model individual-to-individual interactions. In social networks, nodes are people,

or occasionally groups of people, and edges represent relationships between them (Bernard *et al.*, 1991). However, the definition of a ‘relationship’ between individuals may range from merely being in the same school to being friends, which can be determined by measuring amount of time two people spend together over a short period of time.

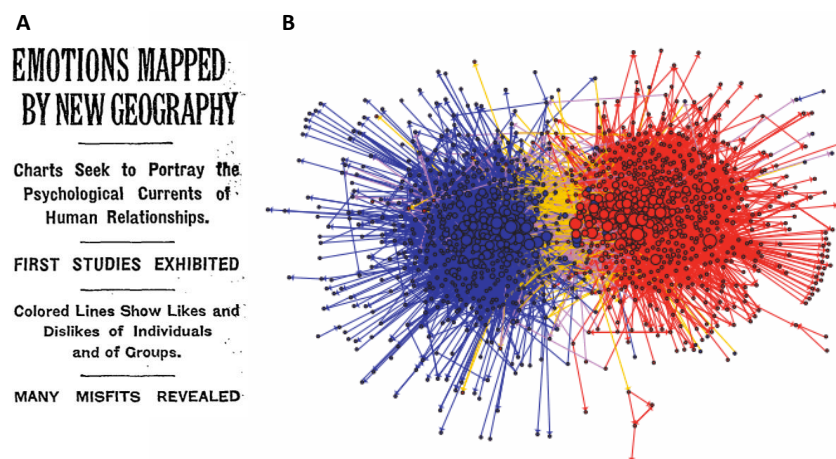


Figure 2.1: **A.** Headline of the NYT report regarding the Moreno’s presentation (Courtesy of The New York Times). **B.** Network of blog-to-blog interaction in 2004 US election. Red nodes represent pro-democrat blogs and blue nodes represent pro-republican blogs. Reproduced from Adamic and Glance (2005)

Thirty years after Moreno’s presentation, in 1963, Stanley Milgram, an experimental psychologist, made an extraordinary contribution to the field by introducing a concept called *six degree of separation*² (Milgram, 1967). He was interested in measuring how many steps (later coined as geodesic distance) it takes someone to reach another randomly selected, unknown individual. He posted packages to 96 randomly chosen individuals in Omaha, Nebraska. Packages contained a fake Harvard passport, his home institution, and a note asking the recipient to re-send the package to the target individual. The target individual was one of the Milgram’s friends based in Boston. The recipients were informed with full name and postal address of Milgram’s friend however, they weren’t allowed to directly post the package to him, but to post the package to one of their close friends. The next recipients were asked to record the sender and receiver in the package and repeat the process until the package reach Milgram’s friend in Boston. One sixth of the packages finally found their way to Milgram’s friend; which, by considering the distance and the randomly selected candidates, was a remarkably high number. Milgram then examined the *path* that each of the packages travelled to get to his friend and concluded that a distance between

²This term was not used by Milgram in his original paper. It was coined after a play at Broadway titled ‘six degree of separation’ inspired by Milgram’s experiment.

two randomly chosen people is approximately six people. Another striking observation about Milgram’s experiment is that although the recipients belonged to tight *communities* in remote areas of Omaha, they nonetheless found their way to the Milgram’s friend in Boston. Six degree of separation was later formulated mathematically as the *small-world* phenomenon (Watts and Strogatz, 1998) in networks, where each node belongs to a segregated community (Omaha vs. Boston), but there are still some shortcuts which guarantee a relatively short geodesic distance between each pair of nodes. Although his conclusion is based on a small number of people and there are controversies about his candidates’ selection criteria, it still motivated other researchers, as well as himself, to continue his work on finding the average shortest path length between nodes of a extraordinarily large network (Bernard *et al.*, 1988; Kleinberg, 2000).

For technical discussion of measuring level of segregation see section 3.1.4

Although the study of social networks initially started by Moreno, it was not limited to Milgram’s experiment or the small-world phenomenon. Today, thanks to the Internet, social networks can be studied over extraordinarily large scales with fewer prefixed assumptions. Nowadays, graph theoretical methods are not only practical tools to investigate one of the most complicated within-species interactions, they are also a plausible mathematical model to tackle challenges such as organised crimes (Krebs, 2002) and human trafficking (Milward and Raab, 2002).

For technical discussion of measuring level of integration see section 3.1.3

Since 1736, there have been enormous advancements in the field of graph theory. These advancements go beyond the application of conventional methods to empirical networks to modelling the behaviour of latent structures hidden in a network by *random graphs*. The first random graph that aimed to model real-life phenomena was introduced by Erdős and Rényi (1959) and independently by Solomonoff and Rapoport (1951). Later, this model was named after these two Hungarian mathematicians. The Erdős-Rényi random graph model (ER) works under the assumption that all nodes of a network have equal probabilities to be connected. This assumption leads to an approximation of a Poisson degree distribution³. This model has been the most widely studied network model since its introduction. However, recent findings demonstrate that it fails to completely replicate the structure of the real-life networks. For example, ER models fail to model the structures which are present in natural systems and, therefore, it can not capture the effect of hubs within a network (Albert and Barabási, 2002).

Decades after the ER model was introduced, in 1998, Watts and Strogatz introduced a model for ‘small-world’ networks (SW) based on Milgram’s six degree of separation model (Watts and Strogatz, 1998). SW models capture the reconciliation of simultaneous segregation and integration of connectivity profiles within a network, which was shown to be the case in several natural phenomena such as affinity groups, seismic networks etc.

For technical description of Small-World phenomenon; see section 3.1.6

³For detailed description of this degree distribution please see section 4.1.1

The degree distribution of SW models sits between a Dirac Delta function (a perfect lattice or regular network) and a Poisson distribution (ER model). Networks with SW structures are more likely to exhibit a collection of highly connected nodes, which facilitate the integration between segregated sub-networks. In some special cases, these hubs tend to exhibit significant connectivities to each other. Such tightly connected sub-networks of hubs have conventionally been referred to as a *Rich Club* (Colizza *et al.*, 2006a). Figure 2.3 illustrates a modular structure and Rich Club organisations on synthetic networks. The SW model was introduced with the aim of modelling the evolution of rewiring between nodes within a network with a fixed number of nodes, and it cannot be considered as a generative model. Moreover, the degree distribution of the SW model is not as homogeneous as the ER model, it still fails to model networks with highly connected nodes (e.g. the Internet, an aviation network).

For technical description of Rich Club organisation; see section 3.1.9

A year after the Watts-Strogats model, Barabasi and Albert introduced a generative model for networks (BA) (Barabási *et al.*, 2009). Degree distribution of BA networks follows a power-law, where a large majority of the nodes have a small number of connections, but very few of the remaining nodes enjoy an extremely large number of connections. The BA model also predicts that the new attachments of preferential nodes are more likely to happen to those highly connected nodes rather than nodes with a lower rate of connectivity. In graph theory terminology, this effect is commonly referred to as ‘richer gets richer’ or a ‘cumulative advantage’. However, the BA model fails to account for segregation that exists in real-life networks such as a group of friends in a social network. Figure 2.2 shows the topological features of these models by synthetic networks.

The aforementioned drawbacks that are common in representing networks using conventional graph theory models motivated network scientists to investigate a more generalised form of random graphs. The ER model emerged from a simple model that considers all nodes to have an equal probability of forming connection, to a sophisticated stochastic model which divides the network into several subsets and allow each subset to have a different rate of connection, however, it still assumes that each of these subsets still follows a certain distribution (Fienberg and Wasserman, 1981). This class of modelling with random graphs, which is known as the *Stochastic Block Models (SBM)*, uses mixture models and an unknown latent variable to estimate the degree distribution of empirical networks. SBMs were shown to be a prominent set of tools for modelling network behaviour, however, they are not free of limitations. One of the fundamental issues in SBMs is the need to find an optimal number of subsets that a network can be divided into. Although, Daudin *et al.* (2004) introduced an autonomous method to obtain the optimal number of blocks for a certain type of SBMs, it still remains an open question for more general applications. SBMs also suffer from computational inefficiency since finding the best fit to a degree distribution

For a technical description of Stochastic Block Models; see section 4.1.1

requires testing the all permutations of the random network. Bayesian (Bernardo *et al.*, 2003) and frequentist (Jaakkola, 2001) variational methods, as well as other methods based on heuristic search (Olhede and Wolfe, 2014), were shown to be useful in order to rectify the computational burden, but they are still far behind the simpler modelling methods, in terms of computational feasibility.

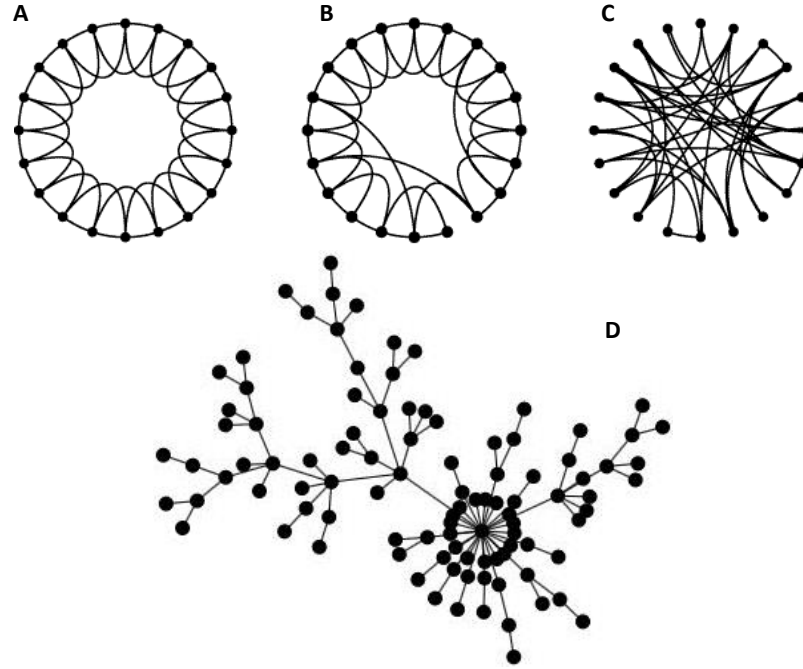


Figure 2.2: **A.** a lattice network with Dirac Delta degree distribution **B.** a smallworld network (SW model). **C.** a random network with Poisson degree distribution (ER model) **D.** a Scale-Free network with power-law degree distribution (BA model). Reproduced from Watts and Strogatz (1998) and Barabási *et al.* (2009).

Applications of Graph theory are not merely limited to modelling real-life phenomena. For instance, it was recently shown that graph theory can also be useful when integrating different types and, consequently, facilitate the statistical inference between them. One of the well-known examples is *multilayer networks* (Kivelä *et al.*, 2014). Multilayer networks are not a new concept in the field of network science, as they were introduced in late 90's. For example, Carven and Wellman (Craven and Wellman, 1973) introduced 'networks of networks' and few years later, in 1987, David Krackhardt (Krackhardt, 1987) demonstrated their application in social networks. However, it is only recently that they have been drawing an enormous amount of attention and have been used to integrate information in several fields of research, such as the modelling of air traffic (Monechi *et al.*, 2015), rather than social sciences. Multilayer networks can be interpreted within a variety

of contexts. For example, in a temporal context, each layer can be a representation of a snap-shot of an event, where the nodes of layer are repetitive across layers, whereas edge dynamics tend to be different from one layer to another. In a spatial context, multilayer networks can represent different versions of interaction between nodes regardless of their chronological orders. Nevertheless, such networks are not free of flaws since the dependencies and coupling between the layers, which are vital for determining model structure, are still unknown. However, recent efforts have been made to combine Multilayer Networks and Stochastic Block Models for investigating the dynamics of network structures (Valles-Catala *et al.*, 2016; Pavlovic *et al.*, 2015).

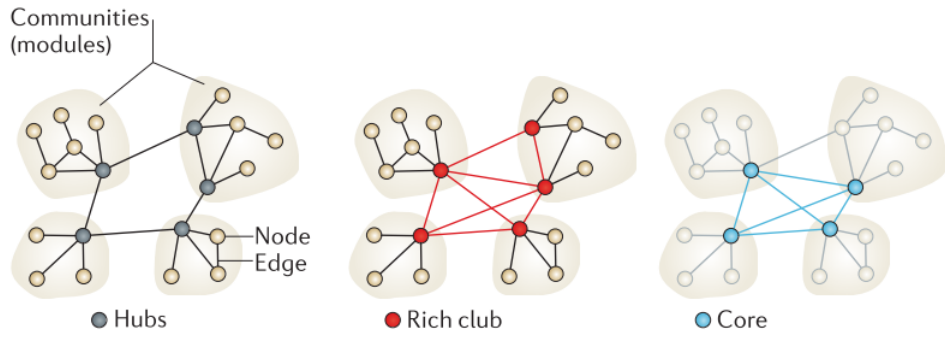


Figure 2.3: Cartoon presentation of a small-world network, with modular structure and a strongly connected core which mediates the modules. Reproduced from Bullmore and Sporns (2012a)

2.3 Network Neuroscience

Within the past two decades, studying the coupling between the function and structure of complex systems has been the focus of several mathematical and engineering disciplines. Examples of such a complex systems can range from human socio-behaviour systems to patterns of spreading rumors and disease. The brain is, also, one such complex system which, despite the enormous advancements in field of neuroscience, has not been fully understood in terms of its structural and functional coupling. Since graph theory was shown to be useful in formulating this relationship, neuroscientists have employed this branch of mathematics to model the connections within the human brain. A wide range of imaging techniques that aim to capture the function and structure of the brain also help neuroscientists gain a better insight regarding temporal and spatial characteristics of the human brain. The first attempts at using network sciences as a theoretical tool to examine the connectivity of the human brain date back to the 90s by (Watts and Strogatz, 1998; Young *et al.*, 1992; Felleman and Van Essen, 1991; Bressler and Ding, 1999). However, describing human brain connectivity as a collection of nodes and edges was coined as *connectomics* in

a paper by Sporns, Tononi and Kotter titled *The Human Connectome: A Structural Description of the Human Brain* (Sporns *et al.*, 2005). Sporns *et al.* argue that describing brain connectivity can be done in three hierarchical spatial levels: micro, meso and macro scales.

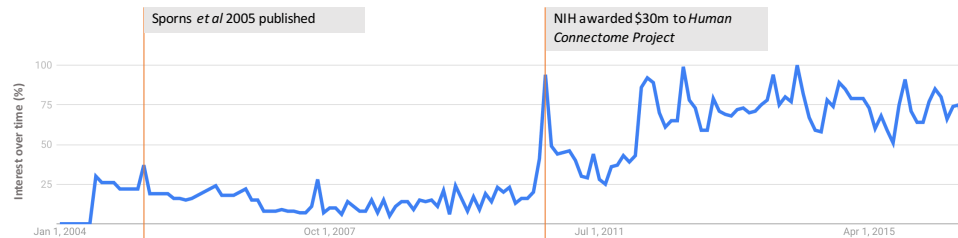


Figure 2.4: Google Trends representing the interest in the topic 'connectome' over a period of 12 years (Jan 2004 - Aug 2016). Spikes (shocks) in the search trend are marked with a brief description.

2.3.1 Microscale Connectome: Axon Terminals and Dendrites

In the context of this thesis, *microscale* is the study of human brain structure in its most finest scale: at the neuron level, by employing electron or light microscopy techniques. Sporns *et al.* argue that this level of description is technically impossible for a human brain with a high number of neurons and synapses and may also be physiologically unnecessary as there is evidence that the functional description of human brain activities cannot be boiled down to a single neuron but to activities of a population of neurons (Stepanyants *et al.*, 2002). The only attempts to map a brain at such a fine scale is the brain of the *Caenorhabditis elegans* worm, which was first published in 1988 (Kenyon, 1988) and it was later improved in 2010 (Félix and Braendle, 2010). The worm's brain was sliced into several ultrathin layers and then electron microscopy techniques were used to identify 6,393 tracks between 302 neurons of the worm's brain. There are several studies about properties of network architectures of the *C.elegans* that confirm that the worm's brain is a complex system, as it shows a balance between network integration and segregation (Latora and Marchiori, 2001). In addition, it was also shown that the worm's brain has highly connected nodes and is consequently a rich club organisation (Towlson *et al.*, 2013).

2.3.2 Mesoscale Connectome: Neurons and Minicolumns

Mesoscale is a higher spatial level of studying the human brain that involves the tracing of axonal projections with different methods of neuroanatomical marking and the micro-analysing of cells and tissues of the brain slices. Axonal tracing methods of mesoscale investigations reveal a subset of a cortical structure called 'mini-columns'. It was shown that, in the human brain, the mini-columns normally contain between 80 to 100 neurons

(Buxhoeveden and Casanova, 2002). Mapping mini-columns is technically infeasible for the human brain, therefore, studying the mesoscale brain is confined to merely mapping scatter regions of the human brain in a very limited number of individuals. However, in 2006, the Allen Institute for Brain Sciences (ABI) released a unique, publicly available mesoscale atlas of the mouse brain (Lein *et al.*, 2007). They used a viral strategy to identify the pathways between different regions of the mouse brain. Each part of the mouse brain was injected with viruses which express green fluorescent protein when a low magnitude current passes through them. The viral injection and current passing through the pathways reveals the minicolumns between different regions (Lein *et al.*, 2007). The mesoscale connectivity map of the mouse brain is extraordinarily valuable as it is known that the mouse and the human brain fundamentally share the same cortical structures, particularly in sensory, motor, emotion and learning/memory system (Lein *et al.*, 2007). ABI also accompanied the mesoscale connectivity map with a gene co-expression map of the corresponding regions which helps researchers link the genomic characterisation of the brain with a close-to-reality map of a brain. Recently published studies about the ABI mouse brain show that it has similar topological properties as the *C.elegans* microscale connectivity map. It has a modular hierarchical structure where modules are tightly connected to each other via a small number of nodes that form a Rich Club organisation (Rubinov *et al.*, 2015). Furthermore, another study suggests that there is a strong link between the pattern of connectivity between 'elite' nodes of the mouse brain and its corresponding gene co-expression pattern of connectivity (Fulcher and Fornito, 2016).

2.3.3 Macroscale Connectome: Regions of Interest and Pathways

A *Macroscale* description of the brain is the highest level of examination of a brain connectivity map. In a macroscale connectome, the brain is divided to several Regions of Interest (ROIs). Although Sporns and colleagues argue that the connectome is the study of the structure of the brain, which he believes to be a derivation of function, the term 'connectome' widens its wing over all other forms of connectivity maps. Today, the functional connectome is as popular as structural connectome as it has shown promising evidence of how the human brain functions. The choice of ROI heavily depends on the modality chosen to record brain activities. For example, the natural resolution of macroscale connectome maps tends to be as large as the number of voxels ($\approx 100k$ nodes), however, it can be decimated to dozens (or hundrands) of nodes according to parcellation schemes. Figure 2.5 illustrates the state of the human brain in macroscale description.

For a brief discussion about of parcellation schemes; see section 5.2.3

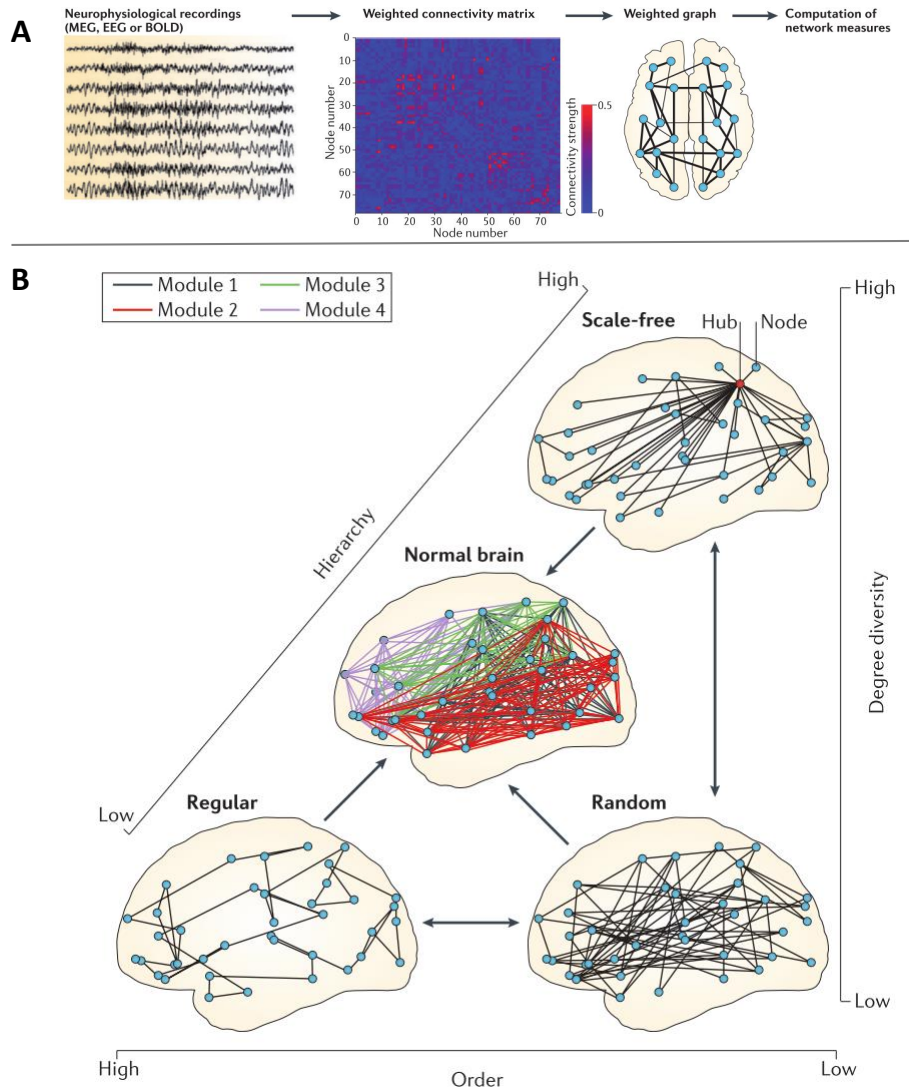


Figure 2.5: **A.** Common procedure for constructing the functional connectome. **B.** Topological description of the human brain, suggesting the human connectome is located between three topological extremes (random, lattice and scale-free). Reproduced from [Stam \(2014\)](#)

These schemes can be derived from prior knowledge of the brain anatomy (anatomical atlas) ([Tzourio-Mazoyer *et al.*, 2002](#)) or data-drive techniques (functional atlas) ([Yeo *et al.*, 2011](#)). On the other hand, the choice of ROIs in Electromagnetic imaging methods (EEG,MEG) is limited by the number of superficial channels (or electrodes) that were used to record the brain activities ([Stam and Reijneveld, 2007](#)). It is also worth noting that, contrary to other connectome scales which are confined to merely a single brain, the macroscale connectome provides researchers with a unique opportunity to study brain connectivity maps across different subjects to examine the effect of different covariates on intra-

subject variabilities.

Analysis of the macroscale functional connectomics of the human brain is one of the connectome areas which draws the most attention. The formation of edges between each pair of regions (or channels) is determined by the measure of temporal dependencies, such as the family of correlation or information theory measures (Smith *et al.*, 2011a). Besides the temporal dependencies for forming an edge between two nodes, more sophisticated methods, mostly inspired from economics, help neuroscientists measure the causation between nodes which results in an directed functional connectome. (Roebroeck *et al.*, 2005) However, due to computational impracticalities producing directed graphs of the brain is limited to dozens of nodes at best (Costa *et al.*, 2015).

One of the very first attempts to use graph theory to describe macroscale functional connectivity maps was done by Stam (2004), who used MEG recordings of five healthy subjects to examine whether the network of functional connectivity follows a small-world model. In his network, nodes were defined as 126 MEG channels and edges were defined as the synchronisation likelihood between all pair-wise combinations of channels. He argued that in all five MEG frequency bands the connectivity networks demonstrate a small-world behaviour. The intensity of the phenomenon, however, varies from one bandwidth to another.

Soon after Stam's paper, several other studies showed consistency across different modalities such as EEG (Micheloyannis *et al.*, 2006) and fMRI (Ferrarini *et al.*, 2009). To date, there has been a lot of effort to unveil other graph theoretical features of the functional connectomics. For example, it is well established that the functional connectome has a hierarchical modular structure (Meunier *et al.*, 2010), which demonstrates a segregated nature of the human brain, but in the meantime, a small number of hubs, which are also highly connected to each other, forms a Rich Club (RC) organisation (Van Den Heuvel and Sporns, 2011). It was shown that the RC organisation acts as a backbone to the functional connectome (van den Heuvel *et al.*, 2012a) and facilitates integration in the brain (Xu *et al.*, 2010). Furthermore, an extraordinary finding from functional connectomics shows that the functional connectome is a *cost-efficient* topology (Vértes *et al.*, 2011; Bassett *et al.*, 2009; Achard and Bullmore, 2007). In other words, it is constrained by a trade-off between minimisation of wiring cost and topology. Surprisingly, the findings about the human brain's macroscale functional connectome are consistent with topological structures that were reported from other species, over different spatial scales. For example, almost every connectome study, regardless of the modality or scale, reports a level of complexity which was raised from a constant battle of integration and segregation tendencies. Additionally, the realisation that the human brain is cost-efficient in its rewiring strategies is consistent with Cajal's paradigm of contiguity (Bullmore and Sporns, 2012a). Recently,

For technical description of cost-efficiency; see section 3.1.7

it was also suggested that the topological features of the human functional connectome, including cost-efficiency, are heritable (Fornito *et al.*, 2011; Sinclair *et al.*, 2015) .

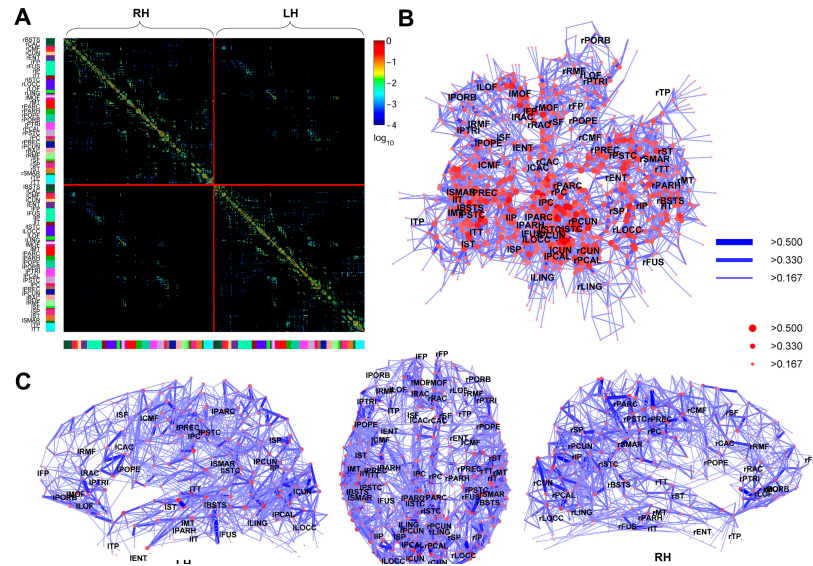


Figure 2.6: Hagmann et al 2008 description of structural connectome: **A.** Adjacency matrix of the structural connectome of healthy brains. **B.** Topological presentation of nodes configured based on Kamada-Kawai force-spring layout **C.** Spatial reconfiguration of figure B with brain region coordinates. Reproduced from Hagmann *et al.* (2008)

Another branch of studying the macroscale connectome is focusing on a structural connectivity map the human brain. Analysis of a structural connectome is based on tractography analysis of the brain using diffusion MRI data (Hagmann *et al.*, 2008), see Figure 2.6 for more details. Besides tractography, other methods such as structural co-variance which is centered on local anatomical features of the human brain, mostly on cortical thickness, were also shown to reveal physiologically important features of the brain topological configuration (Alexander-Bloch *et al.*, 2013). In structural imaging, a parcellation scheme determines the nodes of brain networks, whereas, white matter fibres (in the case of tractography) or similarity between anatomical features across subjects (in the case of structural co-variance) determines the edges. In 2008, Patric Hagmann and colleagues (Hagmann *et al.*, 2008) published the first analysis of the structural connectome in the human brain. They showed that the structural connectomics of the human brain also demonstrate the features consistent across all three spatial levels (micro,meso and macro) of the connectome. Their careful examination of topological and spatial features of the brain structural network led them to conclude that the structure-function integration in the human brain is facilitated through a central core within the human cortex.

2.3.4 Reproducibility and the dead Salmon

In 2008, a poster presented at the annual meeting of the Human Brain Mapping Organisation showed that functional images of a dead Salmon fish suggested a level of activations (Bennett *et al.*, 2009). Although the activation maps of the Salmon's brain were not corrected for multiple comparison error, an established technique to avoid false positives, but it highlights how malpractices in statistical analysis of the brain imaging can result in meaningless results.

Despite the ground breaking findings in network neuroscience it is not yet immune to false findings since it heavily relies on sensitive statistical inferences. There are also several remarkable controversies about the analysis of the human connectome, which make it even more vital to note that network neuroscience is still a young interdisciplinary field which requires exceptional scrutiny in both analysis and interpretation of its results.

The connectome is known to be highly sensitive to pre-processing steps, which are used from the early stages of acquiring the data from the modalities to form the connectivity maps. For example different strategies for regressing out the effects of low-frequency physiological confounds, such as cardiac and respiration processes, can introduce large amounts of negative correlations, however, to date there is no explicit interpretation for the meaning of anti-correlations between regions of the brain. More importantly, other confounds during scanning sessions, such as head movement, can inflate the significant effect in inference between a healthy and a patient group. For example, it was shown that the differences in structural connectome of Autistic patients are due to the uncorrected head movement in the brain, which shows how essential it is to take essential precautionary actions, such as recent motion correction methods (Power *et al.*, 2014) to address this issue.

Recently, it was also shown that the effect of auto-correlated noise in the data remarkably affects the quality of hypothesis tests due to inflation in degree of freedom caused by dependencies between the acquired data-points. Recent studies suggest that using heterogeneous methods to correct the similarity measures (e.g. correlation coefficients) in functional connectome. These issues show how vital it is to carefully inspect the results for different data-sets and under different circumstances. Careful inspection of results ensures that the findings about the human connectome and its alterations under different neuropsychiatric disorders are reproducible enough to be translated to clinical practice in the future. Reproducibility can only be ensured through repeating the experiments on a large number of subjects from different acquisition sites which seems to be economically impractical. Recent extraordinary efforts under 'data-sharing' schemes (Poline *et al.*, 2012) provide a unique opportunity for researchers to examine the reliability of their findings to avoid any over-interpretation.

For deeper discussion about estimating the degree of freedom; see chapter 5

2.3.5 Summary

In this chapter, we gave a brief background regarding the topics, that will be covered in more details, throughout this thesis.

- The history of the imaging techniques have been discussed, as well as the different modalities and the steps required to process the brain activities.
- A brief background regarding the recent advances in network sciences, as a powerful mathematical tool, has been presented
- Connectome, a new field of describing the human brain by using network sciences, has been discussed. A brief description of each scale of connectome was presented.
- Finally, challenges in the field were discussed and the possible solution to mitigate these challenges were mentioned.

Chapter 3

Topological Changes in a Network of the Functional Connectivity of the Human Brain in Schizophrenia

As discussed in 2.2, graph theory has a long-standing history of describing the natural phenomena including the functional and structural connectivity of the human brain. It was also shown to be useful in identifying bio-markers aiming to gain a better understanding of neuro-psychiatric disorders.

One of the most prominent examples of applying network neuroscience methods to identify changes due to neuro-psychiatric disorders is *Schizophrenia*. Prognosis obtained from the investigation of functional and structure connectomes of the patients diagnosed with the disorder suggest a high level of 'dys-connectivity' across the brain, as well as the less prominent role of hubs (Wang *et al.*, 2010; Lynall *et al.*, 2010; Bassett *et al.*, 2010). The term 'dys' in the beginning of the the word connectivity refers to abnormalities (either decrease or increase) in connectivity between brain regions (Stephan *et al.*, 2006; van den Heuvel and Fornito, 2014).

Previous examination of the structural connectome of patients with Schizophrenia reveals an increase in the level of segregation (Zalesky *et al.*, 2011; van den Heuvel *et al.*, 2013a) while showing a decrease in level of integration and global efficiency (Ottet *et al.*, 2013; Van Den Heuvel and Pol, 2010; Zhang *et al.*, 2012; Zalesky *et al.*, 2011). While, the structural connectome of the disorder shows a consistency between different modalities and pre-processing strategies, analysis of the functional connectome suffers from contradictory results across studies and previous findings from the structural connectome. For example, the functional connectome in Schizophrenic brain networks shows a reduction in the level of segregation, but an increase in integration (Alexander-Bloch *et al.*, 2010; Liu *et al.*, 2008;

Becerril *et al.*, 2011; Lynall *et al.*, 2010). However, there are several other studies which suggest the opposite (Yu *et al.*, 2013a; Micheloyannis *et al.*, 2006).

In this chapter, we show how graph theoretical methods, such as measures of integration and segregation, can be used to describe functional relationships between different regions of the human brain. Furthermore, hoping to gain a better understanding about the nature of the human brain, we also show how these mathematical methods can be used to gain a better insight in understanding the effects of Schizophrenia on the functional connectome of the brain. Specifically, motivated by inconsistencies in investigations of the function connectivity in schizophrenia we use a published schizophrenic cohort (Lynall *et al.*, 2010) and use a different parcellation scheme to examine how the results are changed. Eventually, we propose a novel framework for conducting statistical inference between a Rich Club organisation for two groups. This framework is useful for studying intra-subject variability of a Rich Club structure which is commonly ignored by group-averaged networks due to technical difficulties.

3.1 Methods

In this section, we start with describing the approach used to construct the functional connectome of the healthy and schizophrenic subjects. In addition, we discuss the empirical graph theoretical measures that help us interpret topological properties of the subjects' functional connectomes. Throughout this thesis, we frequently use the methods that were described in this section with appropriate cross-reference.

3.1.1 Constructing Network of Functional Connectivity

To construct the functional connectivity networks, brain scans of each subject were parcellated using the AFNI325 atlas into 325 non-overlapping Regions of Interest (ROIs). 28 regions were later discarded due to missing data in some individuals. BOLD timeseries within each ROI were extracted and averaged. This result in only 297 timeseries per subject. We then followed the approach proposed by Achard *et al.* (2006) to form the functional connectivity using wavelet transformations. Time series were decomposed into wavelets of four specific frequency intervals: scale 1, 0.125-0.250 Hz; scale 2, 0.060-0.125 Hz; scale 3, 0.030-0.060 Hz; and scale 4, 0.015-0.030 Hz as suggested in Lynall *et al.* (2010). We then choose scale 2 which was shown to represent much of the brain functional characteristics (Lynall *et al.*, 2010). Eventually, correlations were applied between wavelet coefficients of all possible pairs (Percival and Walden, 2006) which results in a 297×297 adjacency matrix per subject.

We use Density Thresholding (DT) to construct functional networks. A network for each subject is thresholded over a range of densities which starts at 5% and ends at 50% with increments of 5%. The range of density was selected as below 5% the individuals may have isolated nodes and above 50% the individual connectomes may demonstrate random-graph behaviour. In each density step, we only preserve supra-threshold coefficients and set the rest of the edges to zero. For this chapter, we provide a short description about the thresholding method as, later in section 5.1.10, we thoroughly discuss different thresholding methods and their consequences on examining the human connectome.

3.1.2 Network-based statistics

Before starting the investigation of the topological features of the functional networks, it is essential to examine the changes in the correlation coefficients of edges between healthy and schizophrenic subjects. Investigating the correlation coefficients between each pair of nodes helps to avoid an abstract interpretation of the human brain organisation. We use non-parametric hypothesis testing to identify significant changes in correlation coefficients of the functional connectome due to schizophrenia. However, conducting an edge-based mass-univariate testing can be computationally cumbersome and under-powered, since a network with n nodes requires $\frac{n(n-1)}{2}$ hypotheses tests. A large number of hypothesis tests increases the probability of false-positive rates, an effect that is conventionally called Multiple Comparison Error (MCE) (Nichols and Hayasaka, 2003).

To address these issues, we use Network-Based Statistics (NBS) (Zalesky *et al.*, 2010) which is a topological analogue to the cluster-based thresholding of statistical parametric maps in neuroimaging (Hayasaka and Nichols, 2004; Bullmore *et al.*, 1999). This method calculates the t-statistics of each pair-wise connection and then transforms the connectivity maps to a set of components constructed by removing edges with a t-statistic below a given threshold. Each component is constructed using a breadth first search through all possible links. Contrary to the FDR which independently correct the p-values of every single edge, the NBS accounts for false-positive rates of components, instead of edges, and therefore offers a greater power in mass-univariate testings (Zalesky *et al.*, 2010). Eventually, the p-value is calculated by permuting the subjects for five thousand times and recalculating the statistics to form a null distribution. The choice of threshold determines how the method differs from conventional FDR. As recommended in the literature, we examine the NBS technique over a range of thresholds to ensure that the findings are independent of test sensitivity. It is worth noting that, the NBS is not an alternative to FDR. It is rather a complimentary method to boost the execution time as well as the power of the hypothesis testing (Fornito *et al.*, 2013).

The NBS detect significant changes in edge weights between healthy and Schizophrenic

group. We later, investigate how the changes in strength of the edges can lead to further topological changes between the two groups.

3.1.3 Integration Measures

As discussed in section 2.1, anatomical examination of brain regions suggests that the brain regions are known to be highly integrated (Sporns, 2011a). This integration can be mathematically quantified by investigating the paths between two nodes in a topological description of the connectome. In graph theory terminology, a *Path* is a collection of nodes and edges which represent the shortest topological distance between two specific nodes (Bouttier *et al.*, 2003). The average shortest Path between all nodes of a network is called the *characteristic path length* (L) (Watts and Strogatz, 1998). A smaller characteristic path length, suggests a higher level of integration within a network. Throughout this work we use Path Length for Characteristic Path Length, unless otherwise stated. The characteristic path length can be quantified as:

$$L = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N, j \neq i} \sum_{a_{lm} \in g_{i \leftrightarrow j}} a_{lm}}{n - 1} \quad (3.1)$$

where N is number of nodes, a_{uv} is the weight of the edge between node u and v . According to Eq. 3.1 if a node happens to be isolated, its shortest path turns out to be infinity which leads Eq. 3.1 to be infinity too. To address this issue, Latora and Marchiori (2001) proposed a similar method, called global efficiency, which is an average of the inverse shortest path, therefore, the shortest path of isolated nodes is zero.

3.1.4 Segregation Measures

The human brain is also known to be, functionally and structurally, a segregated network which leads the group of nodes to be highly intra-connected rather than inter-connected (Sporns, 2011b). This group of intra-connected nodes leads to the formation of cliques and modules within the network. In this section, we use the metrics of the clustering coefficient and modularity as two commonly used tools for measuring the level of segregation in the human connectome.

Clustering Coefficient and Local Efficiency

Clustering coefficient measures, the prevalence of neighbours of two connected nodes n_i and n_j also happens to be connected with each other (Watts and Strogatz, 1998; Newman, 2003a). In other words, the clustering coefficient is a ratio of the number of complete

triangles around a node over the number of all triangles that can be formed around a node. the number of triangles around node i in a binary network can be obtained as:

$$t_i = \frac{1}{2} \sum_{j,h \in N} a_{ij}a_{ih}a_{jh} \quad (3.2)$$

where h and j are node i 's neighbours. Therefore, the clustering coefficient for node i can be mathematically quantified as:

$$C_i = \frac{1}{n} \sum_{i \in N} \frac{2t_i}{k_i(k_i - 1)} \quad (3.3)$$

where k_i is number of connections to node i . The mean clustering coefficient is the average of clustering coefficient across all available nodes within a network which represents the level of segregation in a network. However, the mean clustering coefficient can be influenced by isolated nodes or leaves in a network (Kaiser, 2008). Although the likelihood of leave and isolated nodes is quite low in a human connectome, to ensure that the mean clustering coefficient is not underestimated, we discard the leave or isolated nodes in mean clustering coefficients.

Similar to the clustering coefficient, Latora and Marchiori (2001) proposed a measure of local segregation, called local efficiency, which can be formulated as (Rubinov and Sporns, 2010):

$$E_i = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j,h \in N, j \neq i} a_{ij}a_{ih}(d_{jh}(N_i))^{-1}}{k_i(k_i - 1)} \quad (3.4)$$

where $d_{jh}(N_i)$ is number of edges of the shortest path between nodes j and h that are only one edge away from node i . It is important to note that although the clustering coefficient and local efficiency are similar, they are not identical as the former measures the probability of formation of a triangle around node i , while the latter measures the topological distance between two neighbours of node i .

Modularity

Another sign of the presence of segregation in a network is modular structures. Group of nodes, which are tightly intra-connected with each other but, at the same time, relatively less inter-connected to the rest of the network are considered as a 'module' (Newman, 2010). Detecting these modular structures in real-life networks is almost impossible without algorithmic methods. Recently, a large number of studies proposed a wider range of techniques to detect the modular structures in a network. These techniques mostly use heuristic methods to find a combination of nodes where the group of nodes has greater intra-connected

edges than inter-connected edges (Fortunato, 2010). In this section, we discuss one of the most well-known methods for detecting non-overlapping modularity detection methods called multi-scale *Modularity Maximisation* method (Newman, 2006). This method measures the level of segregation in a network by two factors. First, module assignments that indicate which module each of the nodes belongs to. Second, the modularity index which, ranging from -1 to 1, indicates to what extent a network is modular. Larger the absolute modularity index, closer the network is to the perfect modular network. The modularity index can be defined as:

$$Q = \sum_{ij} B_{ij} \delta(\sigma_i, \sigma_j) \quad (3.5)$$

where B_{ij} is known as the ‘modular matrix’ and $\delta(\sigma_i, \sigma_j)$ is a delta function which is one whenever module assignment of nodes i and j is the same, and zero otherwise. The most optimal combination of the δ functions is obtained by Louvain method which performs a greedy search through the possibilities to find the one which maximises Q (Blondel *et al.*, 2008). The modularity matrix, B_{ij} , is defined as

$$B_{ij} = f_{ij} - \left[\frac{k_i k_j}{\sum_i k_i} \right] \gamma, \quad (3.6)$$

where f_{ij} is one if an edge connects node i to j (worth mentioning that f_{ij} is binarised form of the a_{ij}), and zero otherwise. The last part of Eq. 3.6 is the probability of an edge existing in an equivalent randomised network which preserves the degree of nodes but dilutes the modular structures. Eventually, a tunable scaler, γ , scales this probability which is called the resolution parameter. When the ‘resolution parameter’ is set to a value smaller than one, it is likely that a_{ij} exceeds its probability which produce larger communities. Conversely, when the resolution parameters is set to a value larger than one, it is less likely that a_{ij} exceeds the probability and this consequently results in a greater number of smaller modules (Betzal *et al.*, 2013).

However, it was shown that maximised modularity is not always scalable by large resolutions as above a certain value the algorithm subdivides the network into singleton modules (Lancichinetti and Fortunato, 2011; Fortunato and Barthelemy, 2007). To find the optimal resolution parameters, we apply modularity detection technique with different resolution parameters which range from $-\log_{10} 10^{-1.5}$ to $-\log_{10} 10^{0.5}$ by increments of $-\log_{10} 10^{-1.3}$ (Betzal *et al.*, 2013). For each increment, we obtain the modularity index for the empirical network as well as a 1k random network. Eventually, we choose the resolution parameter which maximises the difference between the modular structure of the empirical and a random networks.

3.1.5 Network null models

It is important to ensure that the effects which graph theoretical techniques suggest are genuine and are not merely due to randomness. Topological properties of a network are also tightly associated with each other, consequently, a change in one of these features can be reflected in other measures. Therefore, a careful comparison between the empirical results and their null models is essential in order to detect the genuine effects. Null models are randomised graphs which preserve a collection of topological features, while allowing the feature under study to behave randomly (Newman, 2010).

One of the very early null models that are commonly used in the field is the Erdős-Rényi random graph, which preserves the number of nodes and density of the network under study (Newman, 2003b). However, the ER null model assumes that each node has an equal probability of forming a connection with the rest of the network. It has long been suggested that using the ER null model, may inflate the size of any effect as the natural phenomena are far more complicated than what the ER suggests¹ (Hosseini and Kesler, 2013). Alternatively, in this work, we use the Maslov & Sneppen null model (Maslov and Sneppen, 2002). This algorithm randomly picks four different nodes that are connected to each other and swaps the edges in a way that a sub-graph of those selected nodes still has the same degree but none of the new edges already exist in the network. It is worth noting, that by preserving each node's degree, the average degree, and consequently, the overall degree distribution of each null network is identical to the empirical network (for a graphical representation of switching procedure, see Fig 3.1). The switching procedure should be repeated several times, ensuring that the existing structure are completely diluted, however, Maslov & Sneppen suggest that repeating this procedure more than $N \times 4$ time would barely improve the null model (Maslov and Sneppen, 2002).

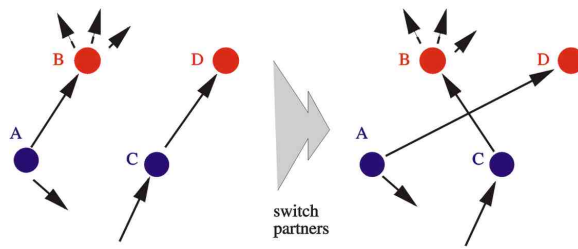


Figure 3.1: Maslov & Sneppen edge swap procedure on four nodes.

¹For a technical discussion about history and pitfalls of ER models see section 2.2

3.1.6 Small-worldness

Following Milgram’s experiment, discussed in section 2.2, the small-world phenomenon appears when nodes of a segregated network are still able to communicate with the rest of the network through short path lengths (Watts and Strogatz, 1998). The reciprocation between segregation and integration in a network commonly suggest the presence of complexity in a network. The structure of complex networks sits between two extremes of lattice and random networks (Watts, 1999). Lattice networks are known to have maximum segregation level, with minimum integration level. Conversely, random networks are known to have a high level of integration, but poor segregation (Newman, 2010).

We measure the level of complexity by comparing the segregation and integration of a network against its null model. Therefore, a small-world network is defined as a network which has an integration level as high as a random graph, but a segregation level remarkably larger than a random graph. To quantify small-worldness, we use a single statistic, σ , which summarise the distance of segregation and integration of the empirical network to the corresponding values in a null model (Humphries and Gurney, 2008), and can be obtained by

$$\sigma = \frac{\Omega}{\Lambda} = \frac{\sum_{j=1}^N L_j^{Empirical} / \sum_{j=1}^N L_j^{Null}}{\sum_{j=1}^N C_j^{Empirical} / \sum_{j=1}^N C_j^{Null}}, \quad (3.7)$$

where Ω is the normalised mean clustering coefficient and Λ is the normalised mean path length. Each of these normalised values is obtained by dividing the empirical clustering coefficient and Path Length by the mean of the corresponding values in one thousand surrogate networks. For σ values remarkably larger than one but Λ values still close to one, the network is known to be a small-world network.

3.1.7 Cost-Efficiency

It was shown that, as the density of a network grows monotonically, the level of integration will also increase as, by assuming that new connections are distributed homogeneously, the new edges facilitate shorter Path lengths between nodes. However, in a specific class of networks, such as financial networks (Vértes *et al.*, 2011) and the human brain (Bullmore and Sporns, 2012a), the association between the increase in the level of integration and density is not linear, and consequently, for a certain density the network demonstrates a high level of integration while its density is still relatively low. This effect is conventionally called ‘cost-efficiency’ and the respective density is known as the cost-efficient density.

Cost-efficiency appears to be a crucial concept as previous anatomical evidence suggest that brain networks are expensive as the brain has spatial constraints, therefore, a new edge should increase the brain’s efficiency in the relatively lowest wiring cost (density)

(Rivera-Alba *et al.*, 2011). Another theoretical implication of wiring-cost is that when the density grows, the network complexity will reduce as the balance between integration and segregation vanishes and the network configuration approaches a fully connected lattice. To determine the cost-efficiency, the level of integration in each density is determined by global efficiency and the wiring cost is measured by the ratio of the number of edges over the total possible number of edges. Therefore, the cost-efficiency can be defined as the difference between the density and the global efficiency and the cost-efficient density is a density where the effect cost-efficiency is maximised (Achard and Bullmore, 2007). Studies suggest that the cost-efficient densities within the brain are most likely to appear in the range of 5% to 30% (Bullmore and Sporns, 2012a).

3.1.8 Centrality Measures

Degree Centrality

One of the simplest measures to investigate the topological features at a local level is degree centrality (DC). This measure calculates the number of edges that are connected to a node. In spite of its simplicity, degree centrality was shown to be a crucial measure in studying the role of each node in the topology of a network (Newman, 2010). This role appears to be crucial, specifically, in small-world networks where degree hubs play a crucial role in connecting a segregated module and cliques to the rest of the network (van den Heuvel and Sporns, 2013). The degree centrality node i can be defined as:

$$k_i = \sum_{j \in N} f_{ij}. \quad (3.8)$$

Betweenness Centrality

Betweenness Centrality (BC) is another tool to quantify the influence of a node in a network. It is defined as the number of shortest paths between two nodes of a network that passes through a certain node (Newman, 2005), as follows:

$$b_i = \frac{1}{(n-1)(n-2)} \sum_{h,j \in N} \frac{\rho_{hj}^i}{\rho_{hj}} \quad (3.9)$$

where ρ_{hj} is total number of shortest paths between node h and j and ρ_{hj}^i is the fraction of shortest paths that passes through node i . Betweenness centrality can be seen as a complementary measure of centrality, since the nodes with high DC are more likely to have higher BC only if they are mediating between a large number of nodes. More simply, nodes with

a high degree and betweenness centrality are those which facilitate the balance between integration and segregation.

3.1.9 Rich Club

Although we discussed the importance of degree hubs in networks earlier, it was shown that the configuration between these hubs can also fundamentally influence the efficiency of the network (Xu *et al.*, 2010). One method of investigating the configuration between degree hubs is Rich Clubs (RC). The RC measures how densely nodes of degree k are connected to other nodes with degree k or higher (Colizza *et al.*, 2006b). In other words, the Rich Club investigates how elite nodes, degree hubs, interact with each other. Rich Club coefficients (Φ) in a network with N nodes and E edges can be quantified by the following equation:

$$\Phi_k = \frac{2E_{\geq k}}{N_{\geq k} \times (N_{\geq k} - 1)}. \quad (3.10)$$

The Rich Club coefficient of degree k , Φ_k , can be obtained by calculating the density of nodes which have a degree of k or larger. However, the coefficient Φ_k is a biased measure for the Rich Club phenomenon since nodes with a higher degree are more likely to be connected to each other merely based on pure chance. Thus, RC coefficients for sorted degrees are likely to monotonically increase (Colizza *et al.*, 2006b). To eliminate this bias, the Rich Club coefficient should be normalised by a relative Rich Club coefficient of their null models (McAuley *et al.*, 2007).

During the generation of null models, the degree centrality of each node should be preserved to ensure that the null networks have the degree ‘clubs’ intact during the randomisation process and merely have their structures diluted. Specifically, the degree distribution of the null model should be identical to the original network. Then, the Normalised Rich Club coefficient, Φ_{Norm} , measures the density of club of nodes with degree k or larger divided by the average density of the respective clubs across null networks, given by:

$$\Phi_k^{Norm} = \frac{\Phi_k^{Empirical}}{\Phi_k^{Null}}. \quad (3.11)$$

The nodes corresponding to Rich Club coefficients with $\Phi_{Norm} > 1$, are considered as a Rich Club (McAuley *et al.*, 2007; Colizza *et al.*, 2006b). We also test each Rich Club coefficient, Φ_k , against its RC coefficients in null networks, Φ_k^{Null} , to ensure that the difference between empirical and the null RC coefficient is significant. In detail, we test whether a RC coefficient is larger than the 95% percentile of the mean null distribution. Any Φ_k with a significantly larger coefficient is considered as part of a RC organisation. Given that the RC curve of observed and null RC will be converged to each other as k decreases and finally

become identical for the lower degree since the degree distribution of the observed and null networks are identical, we avoid the MCE correction as the MCE correction may even make the inference more conservative.

It is worth mentioning that Rich Club analysis provides us with two statistics. One is the RC coefficients, which indicate the level of intra-connection across hubs and the other is RC organisation which indicates which nodes contribute to the setting of significant RC effect. In the next section, we discuss how these two statistics can help us to examine the changes in Rich Clubs of two groups.

Nodal Rich Club Coefficients and Nodal Rich Club Identity

Rich Club analysis is commonly conducted on group-averaged networks, since Rich Club organisation of each subject might be different from other subjects and conducting a between-group statistical inference appears to be impossible and consequently, leads to analysis of RC effect merely on group-averaged networks (Ball *et al.*, 2014; van den Heuvel and Sporns, 2011). Here, we propose two techniques for conducting statistical group inference of both RC organisations and coefficients.

For Rich Club coefficients, we propose the *Nodal Rich Club Coefficient (NRCC)*, which is defined as a matrix of size $S \times N$ where N is number of nodes, S is number of subjects and each element of the matrix is RC coefficient of the club that a node belongs to. For example, if node n has degree of k , the $NRCC(s, n) = \Phi_k^{Norm}$. Forming the NRCC for both healthy and schizophrenia facilitates node-wise statistical inference between individuals' RC coefficients.

Furthermore, we propose the *Rich Club Identity (NRCI)*, which is a binary matrix of the same size and indicate the club membership of nodes. Each element of this matrix is set according to whether node n in subject s has been identified as part of a Rich Club organisation. NRCI can be summarised as below:

$$NRCI(s, n) = \begin{cases} 1 & \text{if } P(\Phi_k \neq \Phi_k^{Null}) < 5\% \\ 0 & \text{otherwise.} \end{cases} \quad (3.12)$$

By calculating the proportion of times a node contributes to a Rich Club organisation across subjects we can form a contingency table and then use the Fisher Exact test to detect the difference between each node across groups. It is worth noting that alternatively, a Chi-squared test can be used for a larger number subjects (Ludbrook, 2008).

Feeder and Local Connections

Besides analysis of the relationships among degree hubs of the human connectome, it is also crucial to examine the interactions between nodes of RC organisation and the rest of the network. These interactions can be classed into two major categories of nodes. First, *Feeder nodes* which are the nodes that are not directly involved in a RC organisation but facilitate the connection between RC organisation and the rest of the network. Second, *Local nodes* which are the nodes that are neither identified as part of the RC organisation nor connected to any nodes within RC organisations.

3.1.10 Disruption Plots

For investigating the local changes across groups, we adopt the Hub Disruption measure, introduced by [Achard *et al.* \(2012\)](#), by combining it with Bland-Altman plot ([Bland and Altman, 1986](#)) to study how local measures of nodes is changed between healthy and schizophrenic subjects. A Bland-Altman plot is a scatter plot, where the x-axis is average degree centrality of each node in control and schizophrenia, and the y-axis is difference between degree centrality of two groups. By fitting a linear line to these point, as Achard et al proposed, we can determine the level of association between changes and the magnitude of graph theoretical measure, this measure is called Disruption Index. It is worth mentioning that nodes that appear on the positive area of the y-axis, are the ones that gained connectivity in patient and, in contrast, the ones within the negative area of the y-axis are identified as nodes that experience a decrease during schizophrenia.

We use disruption plots for a wide range of local graph measures. However, in the case of degree analysis, it is worth noting that as we use a fixed subject-independent threshold for all adjacency networks, the fitted line should always intercept the y-axis in the positive non-zero area. Furthermore, as part of a network experiences increase in connectivity the remaining parts are penalised to hold the balance in density. Therefore, the average changes in disruption plot of degree centrality should always remain zero.

3.2 Data

3.2.1 Sample

To investigate the application of a Rich Block on a human brain network, we used resting-state fMRI data already obtained and analysed using empirical graph theory in [Lynall *et al.* \(2010\)](#). In this study, 15 healthy subjects (age=33.3 years \pm 9.2 years, male=14) and 12 subjects (age=32.8 \pm 9.2, male=10) with chronic schizophrenia were recruited. The patients were diagnosed according to standard operational criteria in the fourth edition of

Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) (Spitzer and Williams, 1980) and were clinically stable during their involvement (exhibiting low symptom ratings and undergoing no change of medication in the preceding four weeks). All were receiving antipsychotic drugs, and four were receiving additional psychotropic medication, but were not treated with their usual medication on the days of scanning to avoid effects on the resultant fMRI data. Healthy volunteers were selected to match the patient group in terms of age, gender, premorbid IQ, years of education and handedness, and screened for major psychiatric disorders using the Mini International Neuropsychiatric Interview (Sheehan *et al.*, 1998). All subjects provided informed consent in writing and the protocol was approved by the Addenbrookes NHS Trust Local Research Ethics Committee.

3.2.2 Resting-State fMRI Acquisition and Preprocessing

To record the brain activities during resting-state, subjects were asked to stay still in the scanner with eyes closed. They were asked to daydream, but not fall into sleep. For resting-state data acquisition, a GE SIGNA system MRI machine operating at 1.5T. For each subject the scan session lasted for 17mins and 12s with TR=2s, TE=40 ms, Flip Angle=70°, voxel size=3 × 3 × 7mm, section skip=0.7mm, matrix size=64 × 64, Field of View (FOV)=240 × 240 × 123mm. The first four slices of each scan were discarded to ensure that the scanner reached equilibrium (resulting in 512 slices per scan). Each dataset was corrected for head movement by realignment and regression and subsequently registered to MNI stereotactic standard space by a 12 parameter affine transform maximizing normalized correlation with a customized EPI template image (within-modality). Registered images were spatially smoothed with a Gaussian kernel (6 mm at full width half-maximum), and the time series were high-pass filtered (cut-off frequency: 1/120 ≈ 0.008 Hz). For more discussion about the pre-processing see section 2.1.

3.3 Results

In this section, we present the results of the methods described earlier on the human brain. For better understanding of the human anatomy and its connection to the results, please use the human brain anatomy presented in Figure A.1.

3.3.1 Changes in Edge Weights During Schizophrenia

In this section, we use NBS to identify changes between unthresholded networks of healthy and schizophrenic functional networks. As discussed in section 3.1.2, the NBS is sensitive to a threshold that is used in t-statistics. Therefore, we repeat the analysis on three different

levels of threshold including 3, 3.5 and 4. Thresholds lower than 3 result in almost all edges to be significantly different and thresholds beyond 4 result in a zero number of edges to be identified as statistically significant.

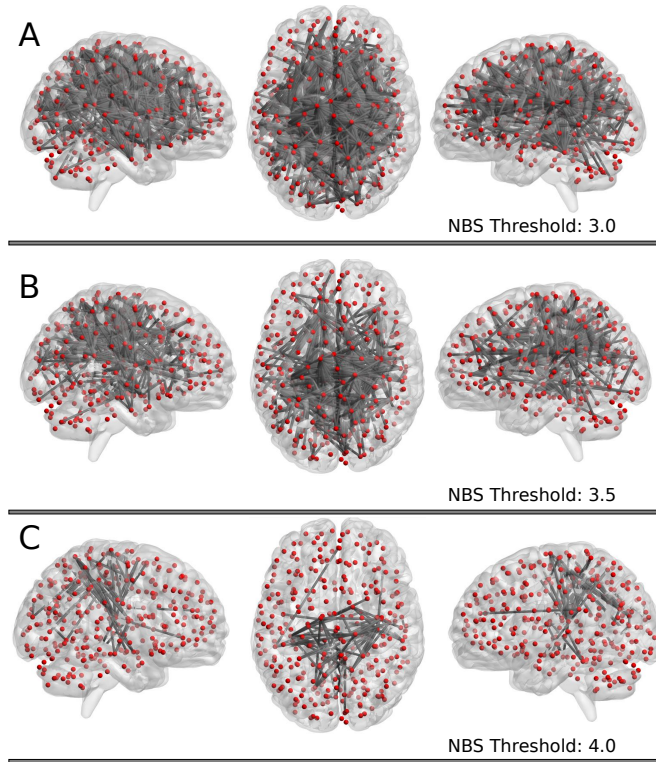


Figure 3.2: Projections of edges that were found to be significantly decreased during schizophrenia by NBS. Results the NBS over different thresholds were shown in each panel.

Although the results of all three levels are shown in Figure 3.2, we only focus on the most conservative test results which are shown in Figure 3.2.C. These results suggest that several edges within the Parietal lobe are significantly decreased during Schizophrenia. Among 41 affected nodes which are shown in Figure 3.2.C, in order, Right Insula, Left Insula, Left Precentral, Left Paracentral, Right Precuneus and Superior Frontal regions are nodes which experience remarkable decreases in their node strength. The most affected nodes that are also connected to these six nodes are mostly within Parietal lobe including the Right Postcentral Left and Right Cuneus, as well as Medial and Middle Frontal Gyrus from Frontal Lobe.

3.3.2 Cost-Efficiency

We start by showing the cost-efficiency across the healthy and schizophrenic subjects. Figure 3.3 shows the global efficiency of each subject, in blue, and confirms that in brain networks the global efficiency grows as the density (wiring cost) increases. In both groups, it is also notable that the association between cost-efficiency and density is not linear as the slope of the growth in global efficiency is greater between 5% to 20%, however, as the figure suggests after $\approx 20\%$, global efficiency demonstrates a monotonic increase with density.

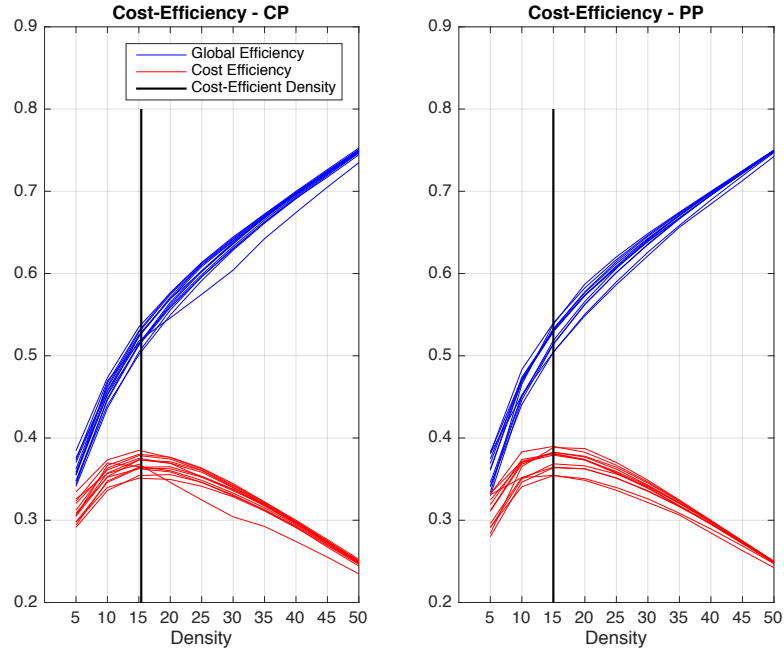


Figure 3.3: Cost efficiency of brain networks in Healthy (Left) and Schizophrenic (Right). Global efficiency of each subject is shown in blue and cost-efficiency is shown in red. The mean cost-efficient density for each group is shown with a vertical black line.

Figure 3.3 also shows the cost-efficiency curve of the brain networks over density. Cost-efficiency, as discussed in section 3.1.7, is a measure of the difference between global efficiency and density. As expected, the cost-efficiency peaks on densities close to 15% and as the slope of the global-efficiency slows down after this density, the cost-efficiency also decreases until the density is $\approx 50\%$ where it reaches its lowest level. It is worth noting that the cost-efficiency of subjects between two groups fail to reject the hypothesis test, and therefore, there the mean cost efficiency between groups is not different.

3.3.3 Global Measures: Integration, Segregation and Complexity

In this section, we start by examining the Path Length as a measure of integration in brain networks and its difference across groups of healthy and schizophrenic patients. Figure 3.4 suggests that Path Lengths are relatively larger for lower densities since in sparse networks, it is more likely that the shortest path between two nodes to be relatively large. Conversely, in highly connected networks (larger densities) it is more likely that two nodes are directly connected via a lower number of nodes and edges. In such networks, the level of integration is higher as the path lengths are relatively low. Complimentary to path length, we also measure the distance between the path length of each subject and its corresponding value in a randomised network. Therefore, values close to one suggest higher similarity between the empirical and randomised Path Lengths. The first row of Figure 3.4 suggests that the both non-normalised and normalised Path Lengths are identical across groups.

The results in the second row of Figure 3.4, suggest that in lower densities, the clustering coefficient is relatively low. A lower level of clustering coefficient means that networks exhibit a lower level of segregation. However, as the density grows, the clustering coefficient also increases. In higher densities, the networks are highly connected and it is more likely that three nodes form a triangle around each other. In addition to the clustering coefficient, we also show that the normalised clustering coefficient, Ω , which is analogous to the distance between the mean clustering coefficient of an empirical network and mean Clustering coefficients of the randomised networks. Normalised clustering coefficients, suggest a higher value in sparse networks as the randomised networks have greater freedom in each realisation to dilute the segregated structures of the networks. As the density grows, the normalised clustering coefficient approaches one which suggests that the level of segregation in empirical networks is similar to randomised networks for higher densities.

For non-normalised clustering coefficients, the changes between the clustering coefficient of healthy and schizophrenic brain networks is not consistent over density. For low densities (5% to 20%) there is no significant difference between the two groups, however, above 20%, the Clustering coefficients in healthy brain networks suggest a higher value. Conversely, the normalised clustering coefficients suggest higher values in Schizophrenic brain networks which is statistically significant across the majority of densities.

Another measure of the segregation which we discuss here is the modularity index Q . We suffice to give a brief discussion of the modularity index, as a global measure of segregation, in this section, however, an in depth discussion regarding the difference between modular structures of two groups is presented in section 3.3.6. The Modularity index, similar to other segregation measures, shows a high value on sparse densities, however, as the density grows the modularity index is decreased. For highly, or fully, connected networks the modularity index is expected to be approximately zero as it seems to be impossible to find a

modular structure in a fully connected lattice. It is notable that statistical inferences failed to provide any significant differences between modular structure of healthy and schizophrenic brain networks.

Eventually, we investigate the small-worldness as a measure of complexity. Results of small-worldness across groups suggest that the measure has larger values on sparse networks. However, as the density grows the small-worldness is decreased until it settles to one where the segregation and integration of the empirical network is equal to the randomised network. It is worth noting, that one important condition for a network to be small-world is that its normalised path length stays close to one which is not the case for very sparse densities. For instance, for a 5% density, the normalised path length is ≈ 1.4 . Therefore, it is unlikely that the network with such a sparse density truly exhibits small-world behaviour. However, the slope of decrease in the normalised path length is greater close to the cost-efficient density, the normalised path length across both groups reaches ≈ 1.15 .

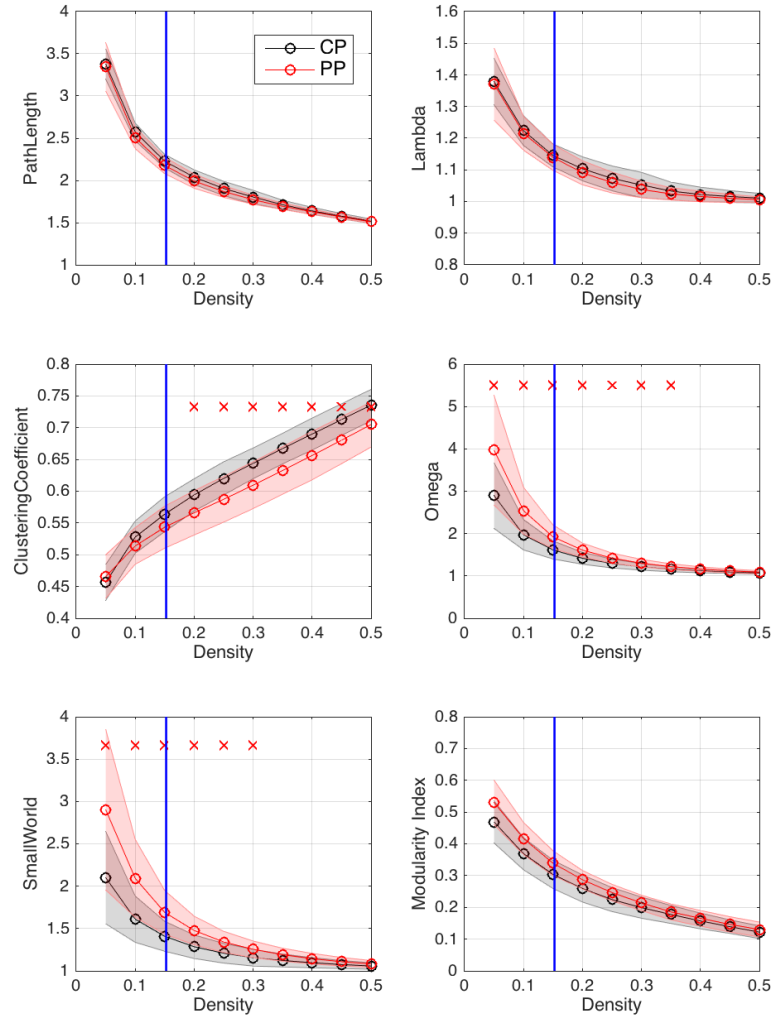


Figure 3.4: Global measures of segregation, integration and complexity. Panel A. shows the global measures for the level of integration: path length (Left) and normalised path length (Right). Panel B. shows global measures for segregation: clustering coefficient (Left) and normalised clustering coefficient (Right) and maximised modularity index (Bottom) Panel C. shows the small-world index as a measure of complexity

Results of the small-worldness curve suggest that brain networks of Schizophrenic patients are more likely to exhibit small-world behaviour since for the majority of densities, including the cost-efficient density, the small-world statistics are significantly larger in healthy subjects. These differences between the small-world index may have been caused by the difference, previously discussed, in the normalised clustering coefficients of healthy and Schizophrenia subjects which later were reflected in the results of the small-worldness.

3.3.4 Centrality Measures in Healthy and Schizophrenia

In this section we investigate the centrality measures in brain networks of both healthy and schizophrenic subjects. To begin with we calculated degree centrality of each node across healthy subjects and then averaged the scores across subjects of each group. Figure 3.5 shows the top 10% of highly connected nodes in healthy subjects. Averaged degree centrality of healthy subjects suggests that the bilateral Precunues and Cingulates dominantly exhibit a high level of connectivity. These areas are followed by slightly less connected nodes such as the bilateral Superior Temporal and Middle Temporal regions.

By preserving the orders similar to degree centrality, we also show the betweenness centrality of each of the degree hubs. It is notable that very highly connected such as the Right Precuneus and Left Cingulate are still among the nodes with high betweenness centrality. However, Right Declive is also shown to be one of the nodes with highest BC value. These two regions are followed by other parts of the Right Precuneus, Right Inferior Frontol, Cerebellum and Left Middle Temporal regions.

We used a similar analysis to investigate the centrality measures in brain networks of Schizophrenic patients. Figure 3.6, top panel, shows the top 10% of highly connected nodes in diseased brain networks. The results suggest that the Left Declive and Right Precunues jointly are the most connected nodes. These two regions are followed by other parts of the Parietal lobe such as the bilateral Precuneus and Cingulate. It is also notable that some parts of the Cerebellum such as the bilateral Declive and Culmen also appears to be highly connected in schizophrenic brain networks.

Analysis of betweenness centrality in schizophrenic brain networks also suggests that among the degree hubs in schizophrenia, the Right Superior Temporal regions as well as the Left Declive and Right Precuneus have the highest BC values. These regions are followed by bilateral Fusifroms and other parts of the Percuneus regions.

Eventually, we examined the proportion of the degree hubs that are similar across two groups. The results suggest that 65% of nodes that were identified as degree hubs in healthy subjects were found to be degree hubs in Schizophrenic patients too. The remaining 35% of nodes that experienced changes are among regions from Partietal lobe including the Cuneus, parts of the Precunues and other regions such as the Thalamus. These regions were substituted by regions from the Frontal lobe such as the Midial Frontal and Inferior Frontal regions as well as the Right Culmen, Left Caudate and Cerebellar Tonsil.

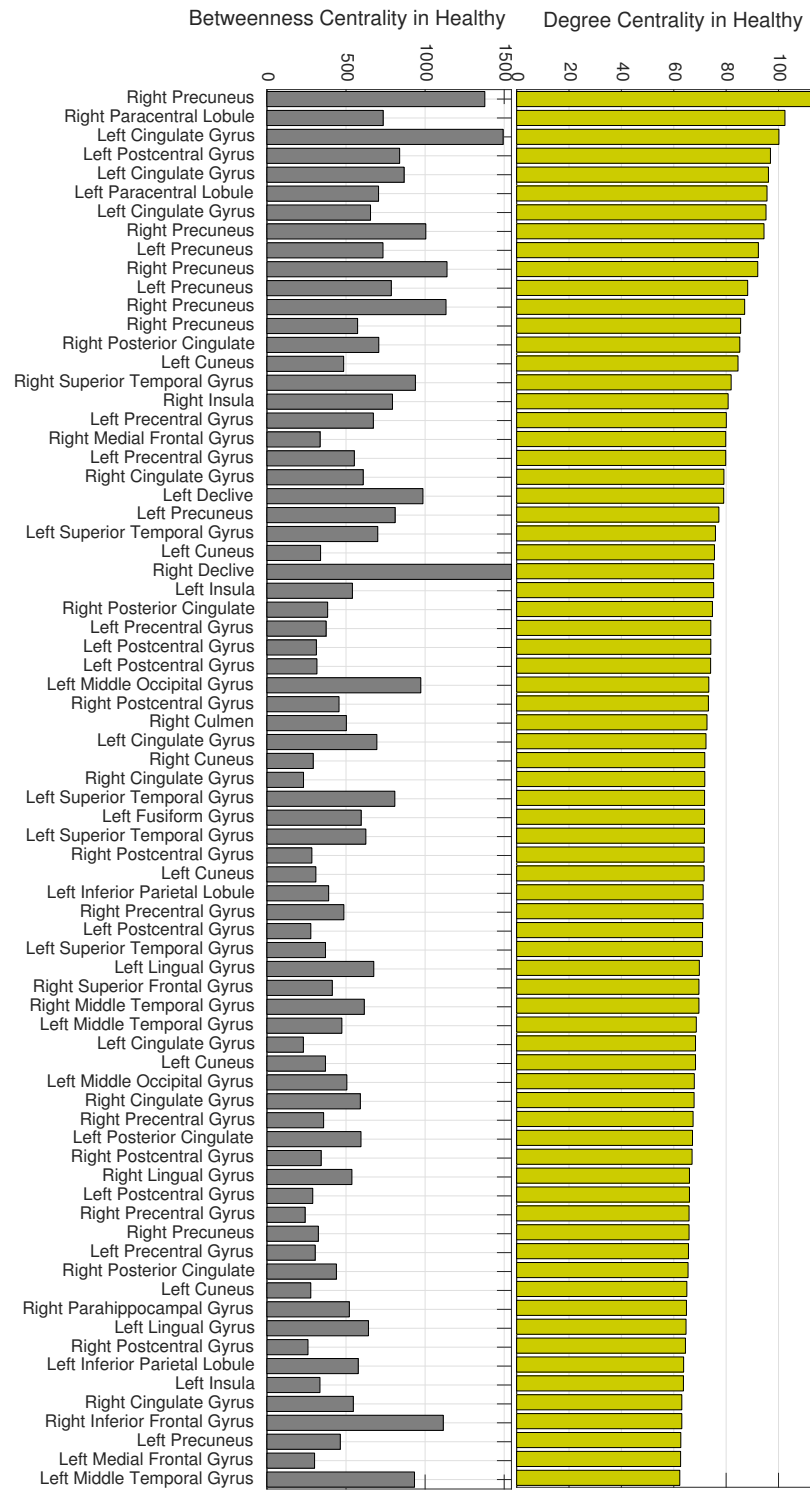


Figure 3.5: Degree (top) and betweenness (bottom) centrality for averaged across healthy subjects.

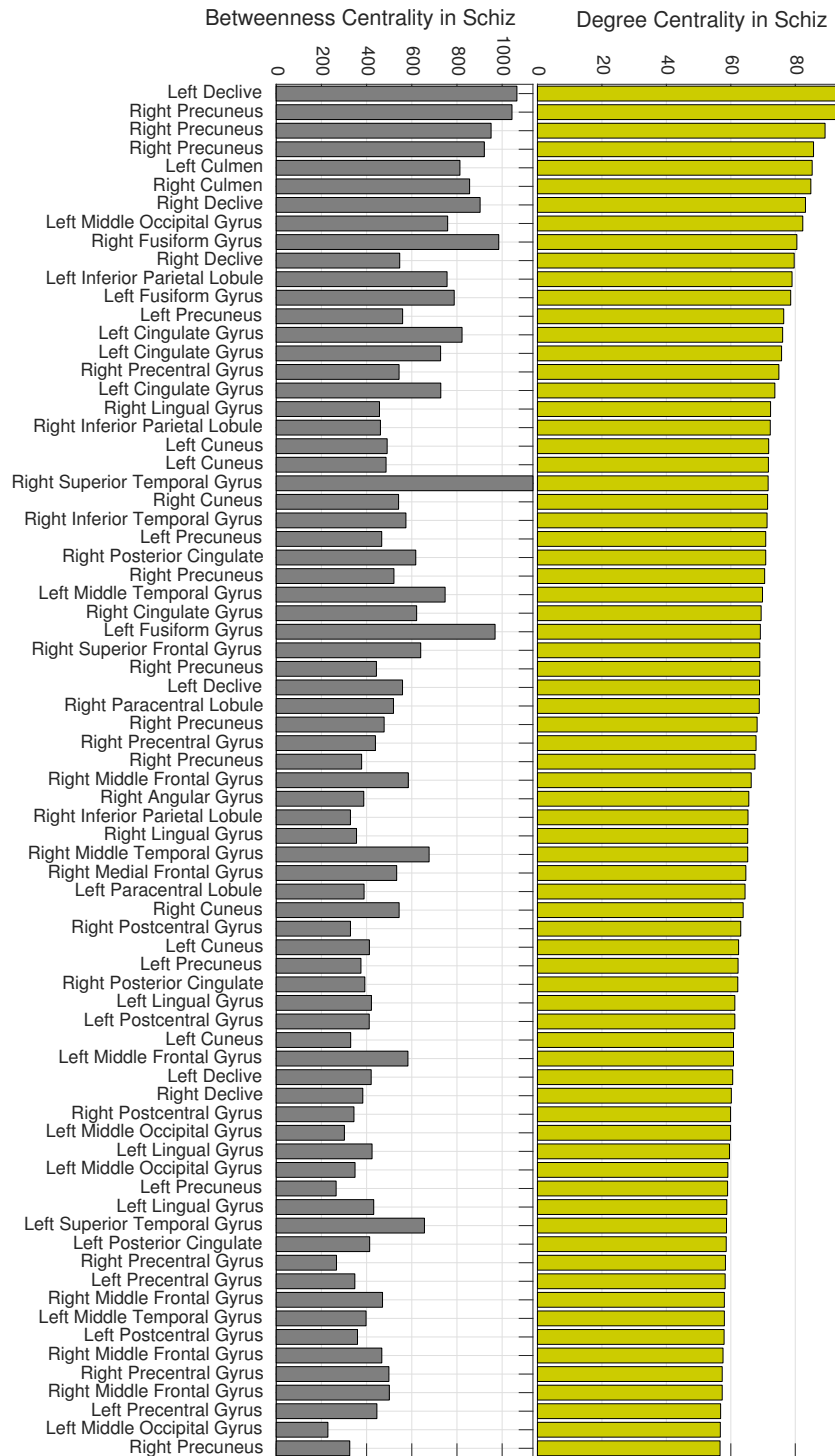


Figure 3.6: Degree (top) and betweenness (bottom) centrality for averaged across Schiz subjects.

3.3.5 Disruption plots of local centrality measures

In this section we discuss the changes in local measures between healthy and schizophrenia subjects. These local measures include local segregation, centrality measures and Nodal Rich Club coefficients. We use disruption plots discussed in section 3.1.10 to track the changes between two groups.

We start with degree centrality which, broadly speaking, suggests a pattern in changes of degree centrality during Schizophrenia since the fitted line to changes shows a slope of ≈ -0.27 which indicates an association between the degree centrality of each node and the amount of change during schizophrenia. In other words, the results of the disruption plots of degree centrality suggest that more prominent degree hubs are more likely to be affected by Schizophrenia. Although this decrease in degree centrality is later compensated by an increase in periphery nodes. The detection of this trade-off between hubs and periphery nodes during schizophrenia is facilitated by the same-density thresholding method which means that the mean changes between healthy and schizophrenia subjects should always remain zero.

We follow by investigating the changes in degree centrality by examining the alteration in by the DC of hub and peripheral nodes. The disruption plot of degree centrality in Figure 3.7 suggests that the most affected degree hubs are the Right Precuneus, Right Paracentral and Left Cingulate regions. On the other hand, the loss in degree centrality of these regions was later compensated by a significant increase in peripheral nodes such as parts of the Cerebellum including Culmen, Pyramis, Tuber and Ivula regions. Also parts of the Frontal lobe including Superior Frontal, Middle Frontal and Parietal Lobes are among the regions which experience a significant increase in their degree centrality.

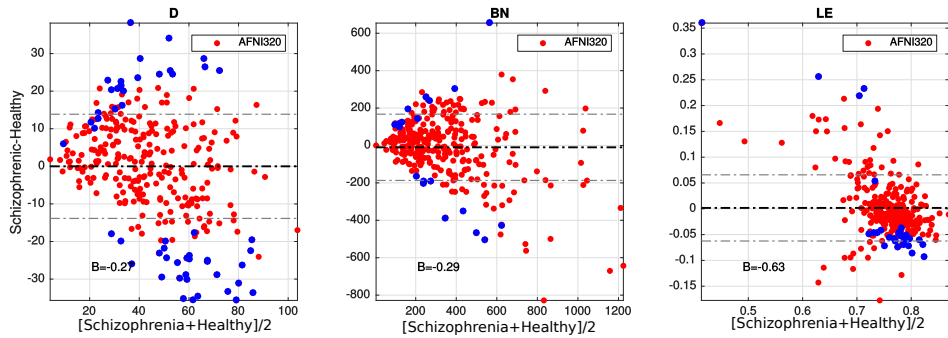


Figure 3.7: Disruption plots of Degree centrality (Left), Betweenness centrality (Middle) and Local Efficiency (Right)

Betweenness centrality is another measure of centrality that we examine here in local level. The slope of the fitted line to the changes, ≈ -0.29 , suggest that there is

an association between the magnitude of betweenness centrality and their changes during Schizophrenia. The average of the significant changes is almost equal to zero suggesting that loss of betweenness centrality in nodes is later compensated by an increase in BC values of another nodes. At the nodal level, the Left Precuneus, Right Insula, Left Postcentral and left Thalamus are among the nodes which exhibit a high level of Betweenness Centrality in Healthy subjects but later experienced a decrease during Schizophrenia. Moreover, nodes that can be identified as nodes with relatively low BC values that experience a significant increase during Schizophrenia includes the Right Superior Parietal, Right Anterior Cingulate and bilateral Superior Frontal regions.

Besides centrality measures, we also examine the effect of schizophrenia on local segregation of brain networks measured by local efficiency. The disruption plot of Local efficiency in Figure 3.7 suggest sthat there is a strong association, ≈ -0.63 , between the changes in local efficiency and the original local efficiency of the nodes. In other words, if a node has a large local efficiency it is more likely to experience a decrease due to schizophrenia. The nodes with high local efficiency that experience a significant decrease includes the Right Postcentral, Left Middle Frontal, Right Cuneus and Right Postcentral. On the other hand, a node that experiences a remarkable increase in its local efficiency is the Cerebellar Tonsil.

3.3.6 Changes in Modular Structure during Schizophrenia

In this section, we show how the modular structure of the human brain is changed during schizophrenia. The results that we discuss in this section were obtained from group-averaged brain networks. For each group, the group-averaged network then thresholded on the cost-efficient density, 15%, which was obtained by analysis of individual cost-efficiency.

As Figure 3.8 shows, the difference between the modularity index of empirical and randomised networks maximises at ≈ 0.93 for both the healthy and schizophrenia group. Therefore, it suffice to examine the brain networks on this scale. Previously it was shown that the average modularity index of healthy and schizophrenia subjects is fairly close (see Figure 3.4), but the group-averaged results of modularity, Figure 3.8, suggest that modularity indexes are 0.36, 0.41 for healthy and schizophrenia subjects, respectively. In addition to the modularity index, the group-averaged brain network in schizophrenia suggest six modules, whereas, the group-average network of the healthy group suggest four modules. It is notable that the number of nodes within each of the modules in the healthy network is almost equally distributed, whereas, the modular structures in schizophrenia suggest that the size of each module can range from 10 nodes (module 1) to more than 96 nodes (module 2).

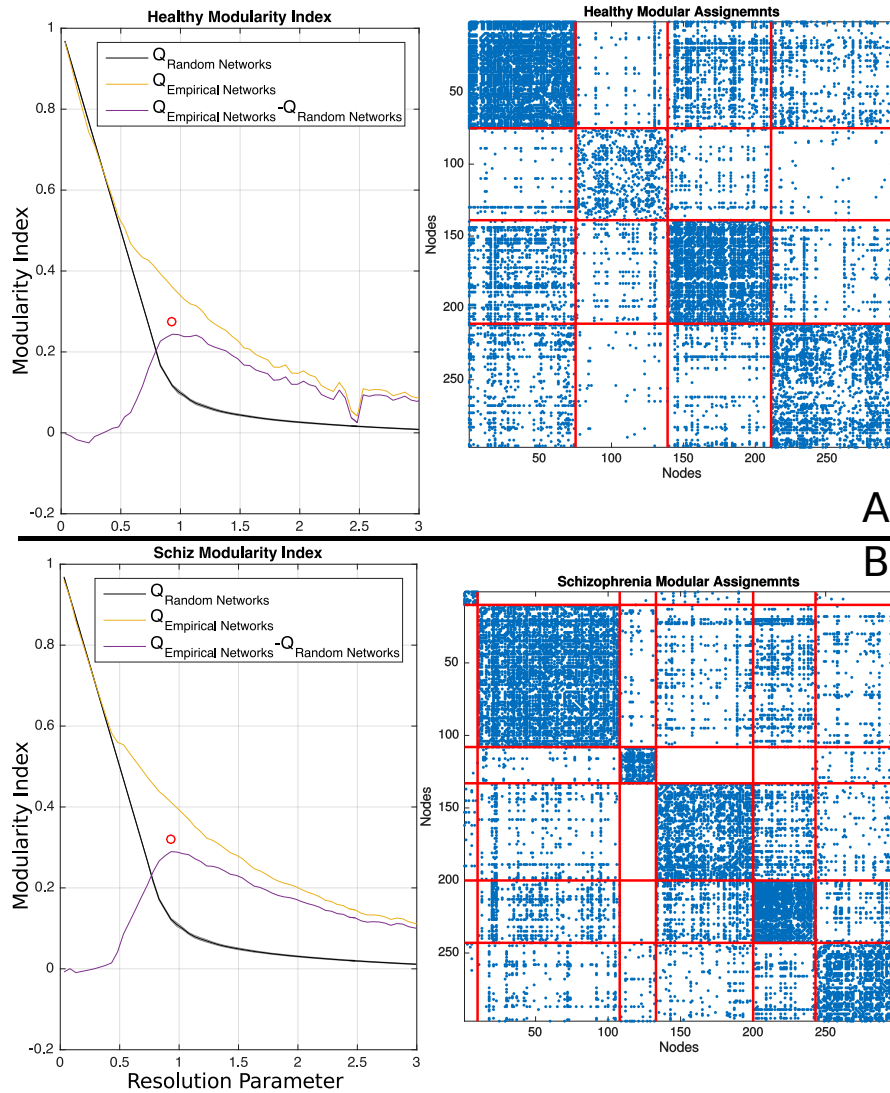


Figure 3.8: **A.** Modularity index across different scales for both empirical and randomised networks in healthy (Left). Module configurations on re-arranged group-averaged network of healthy (Right). **B.** Modularity index across different scales for both empirical and randomised networks in schizophrenia (Left). Module configurations on re-arranged group-averaged network of schizophrenia (Right).

Visual inspection of the module in Figure 3.9 suggests that, although, the modularity algorithm was used on BOLD signals of the brain, the modules are spatially consistent. For example, the majority of nodes that were already known to be part of the Cerebrum were clustered as one module (see module 4; cyan). This effect is even more obvious in the Frontal (module 1; red), Temporal (module 2; green) and Occipital/Cerebellum (module 3) cortices. On the other hand, in schizophrenia, two extra modules, in the Frontal and

Temporal, dilute the spatial consistency across modules.

In terms of module assignments in the healthy brain network, the majority of nodes that formed *Module 1* are from the anterior regions such as the Paracentral, Middle Temporal, Superior Parietal, bilateral Insula, Postcentral and Inferior Frontal. *Module 2* is mainly formed by a combination of regions from the Cerebellum and Temporal lobe including: Tuber, Thalamus, Culmen, Uvula, Fusiform, Parahippocampal, Middle Temporal, Inferior Temporal and Superior Temporal. *Module 3* is mainly formed from the Occipital lobe, Cerebellum and Parietal Lobe which includes Lingual gyrus, Declive, Middle Occipital, Inferior Occipital and Pyramis. Eventually, *Module 4* is formed from regions mostly from the Frontal (middle, inferior, superior) and Parietal lobe which includes the Precuneus, angular gyrus, Thalamus, Caudate and Cingulate gyrus.

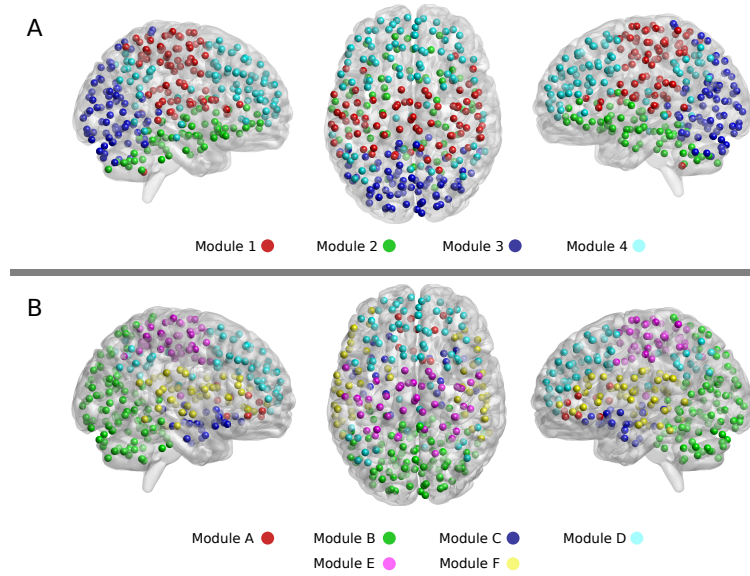


Figure 3.9: Projection of individual nodes and their module assignments on a transparent surface of the human brain.

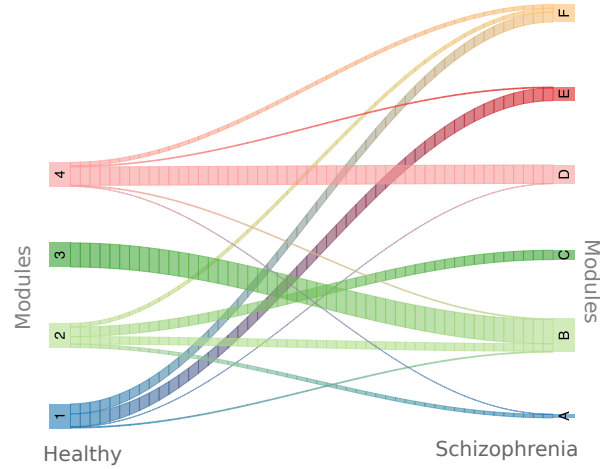


Figure 3.10: A river-plot shows the overlap between modular structure of healthy and schizophrenia subjects.

Examining the modular structure of the functional connectome in schizophrenia suggests that *module A*, which is the smallest module with only 10 nodes, is mainly formed by regions representing the Anterior Cingulate and bilateral Caudate. *Module B* is mainly formed by nodes in module 3 and module 2 of the healthy networks including parts of the Occipital lobe, Temporal lobe and a small fraction of superior Parietal lobe including Middle Occipital, Inferior Occipital, Culmen, Nodule and a small portion of Precuneus. *Module C* is formed by small fraction of module 2 in healthy networks which is mostly located in the Temporal Lobe including the Parahipocampal gyrus, Thalamus and Insula. *Module D* has a large overlap with module 4 of the healthy network as it is mainly formed by the Superior, Medial and Inferior Temporal Lobe as well as anterior parts of Cingulate gyrus. *Module E* mostly overlaps with module 1 of the healthy networks. It mainly represents regions from the Frontal and Parietal lobes such as Precuneus, Precentral and Postcentral lobes. Eventually, module F which overlaps with three modules, (1,2,4) of the healthy networks, covers the Parietal and Temporal Lobe and includes regions such as the Insula, Middle Temporal and Frontal, Inferior Parietal and Lentiform nucleus.

Using Normalised Mutual Information, we approximated the level of association between module assignments of healthy and schizophrenic brain networks and the results suggest that there is more than 59% of the nodes that overlap despite the difference in the modularity index and the block counts.

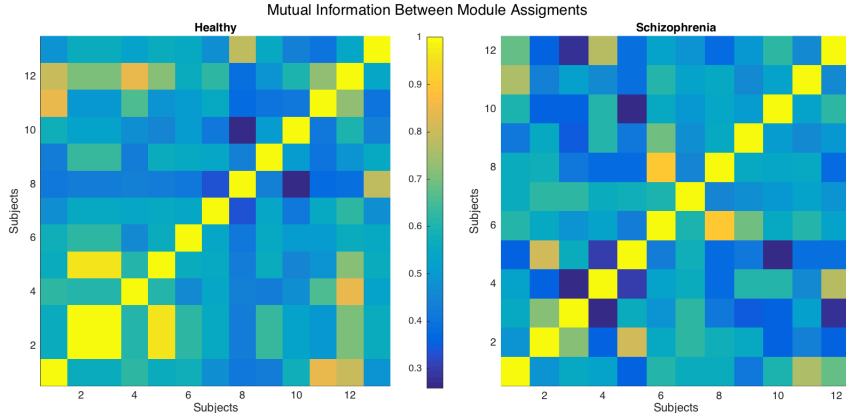


Figure 3.11: Similarity between node assignments of modularity algorithm across subjects of healthy (Left) and schizophrenia (Right)

Using the same technique, we also calculated the level of within group similarity between module assignments of each group. Figure 3.11 shows the NMI score between each pair of the subjects. Hypothesis testing suggests that the mean NMI scores are not significantly different between healthy and schizophrenia subjects.

3.3.7 Changes in Rich Club Organization in Schizophrenia

In this section, we firstly show the Rich Club organisations within the human connectome. We follow the analysis by investigating the difference in the Rich Club coefficients and organisations during schizophrenia. We calculate the Rich Club coefficient for each subject and by using the hypothesis testing we identify the nodes that contribute to the Rich Club organisation of each subject. From a preliminary investigation of the results across subjects, we merely consider nodes which were identified as the Rich Club nodes in more than 50% of the subjects of a group as the Rich Club nodes. However, a thorough framework for conducting statistical inference between the Rich Club coefficients and organisation between groups will be shown later in this section.

Figure 3.12 shows surface projection of the Rich Club nodes of healthy (green), schizophrenia (blue) and those mutually shared between two groups (cyan). Broadly speaking, the Rich Club nodes across both groups are distributed across modules. For example, Table 3.1 suggests that to what extent the Rich Club nodes in healthy subjects has contributed to the modular structures. However, module 1 and Module 3, which represent major parts of the Parietal and Occipital lobes, enjoy a larger proportion of RC nodes in comparison to the remaining modules which contain nodes of the Temporal, Frontal and Temporal lobes.

	Module 1	Module 2	Module 3	Module 4
RC Contribution (%)	56	1.5	44	9

Table 3.1: Porportion of nodes in each module that contributes to the RC organisation of healthy subjects

In schizophrenia, the RC contribution to modules is not as consistent as healthy subjects since only four modules out of six modules show a level of contribution from RC nodes. As Table 3.2 suggests, Module B and Module E suggest that the highest level of RC contribution by 40% and 32% respectively. These two modules are mainly related the Occipital, Cerebellum and Parietal lobes of the human brain. On the other hand, Module A and Module C suggest a zero RC contributions. These two modules are mainly formed by a fraction of nodes from module 2 of healthy subjects and mostly cover the Inferior Frontal and Temporal lobes.

	Module A	Module B	Module C	Module D	Module E	Module F
RC Contribution (%)	0	40	0	11	32	7

Table 3.2: Porportion of nodes in each module that contributes to the RC organisation of Schizophrenic subjects

Bearing in mind that the Rich Club organisations are formed by a collection of highly connected degree hubs and the realisation that they are spatially distributed over the brain, it is essential to examine local measures of RC nodes as well as their association to other parts of the network. For this purpose, we examine the local centrality and efficiency measures of the RC, Feeder and local connections.

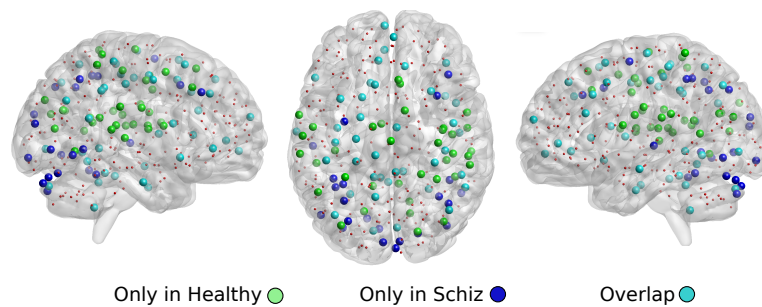


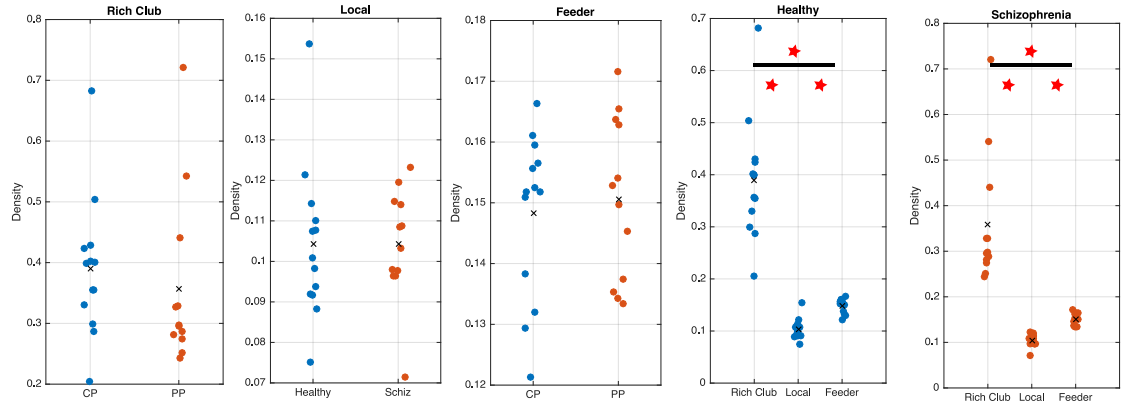
Figure 3.12: Projection of Rich Club organisation in healthy (green), schizophrenia (blue) and overlaps between two groups (cyan)

Figure 3.13.A shows that, in Healthy and Schizophrenia subjects, the density of

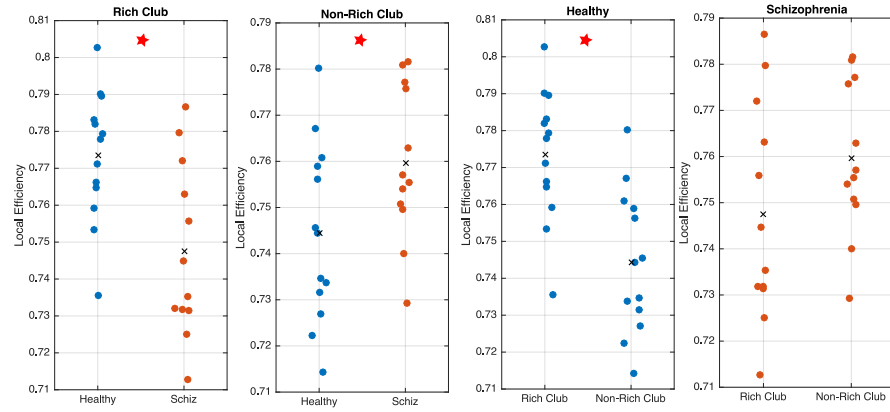
connections between RC nodes is significantly higher in comparison to the density of the feeder and local nodes. However, it is notable that the density of each of these three classes of nodes is not significantly different across groups. Figure 3.13.B shows the analysis of local efficiency of the RC and non-RC nodes. Results suggest that the RC nodes have relatively higher local efficiency than the non-RC nodes in Healthy subjects, however, this effect is no longer the case in Schizophrenic brain networks. The between group analysis of local efficiency of the RC and non-RC nodes also suggests that the RC nodes of healthy subjects have higher local efficiency compared to schizophrenic brain networks. Whereas, analysis of the non-RC nodes suggest that these nodes exhibit higher local efficiency in Schizophrenia. Eventually, Figure 3.13.C shows that Betweenness centrality of RC nodes is significantly higher than the non-RC nodes across both groups. However, inter-group comparison of Betweenness centrality between the RC and non-RC nodes suggests no difference between two classes of nodes.

In addition, we also compare the NRCC and NRCI of healthy and schizophrenic aiming to identify the changes in coefficients and organisation of the Rich Club during the disease. Figure 3.14 shows the NRCC of healthy (top) and schizophrenic (bottom) brain networks. The results suggest that RC coefficients of nodes of the Parietal lobe including parts of the Inferior Parietal, Postcentral, Cingulate, Precunues and Anterior Cingulate regions as well as parts of the Insular region experienced a decrease during schizophrenia. Whereas, regions mainly from the Cerebellum including Pyramis, Declive, Culmen, Uvula experience an increase during schizophrenia.

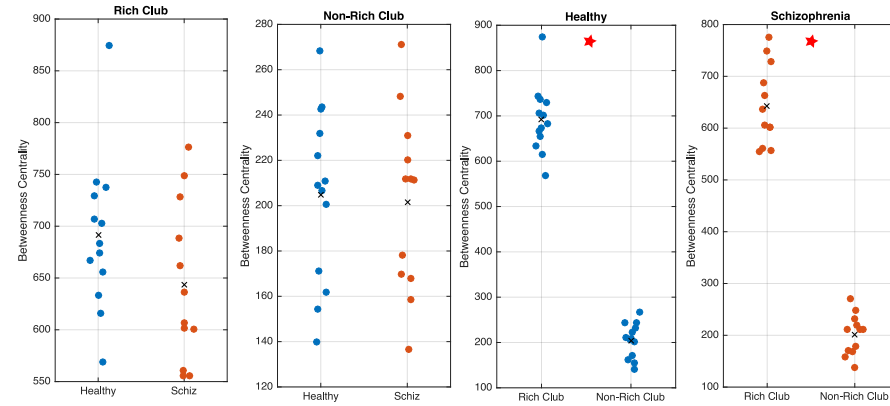
Although we showed how schizophrenia can affect the RC coefficient, it is essential to investigate the changes in configuration of RC organisation during schizophrenia. Figure 3.15 shows the Rich Club organisations of healthy (top) and schizophrenic (bottom) brain networks.



A



B



C

Figure 3.13: Interaction between RC organisation and rest of the network via feeder and local connections. **A.** Analysis of density between feeder, local and RC nodes as well as between healthy and schizophrenia subjects. **B.** Analysis of Local efficiency of local, feeder and RC nodes as well as between two groups. **C.** Analysis of Betweenness Centrality of feeder, local and RC nodes. Red asterisks indicate significant differences.

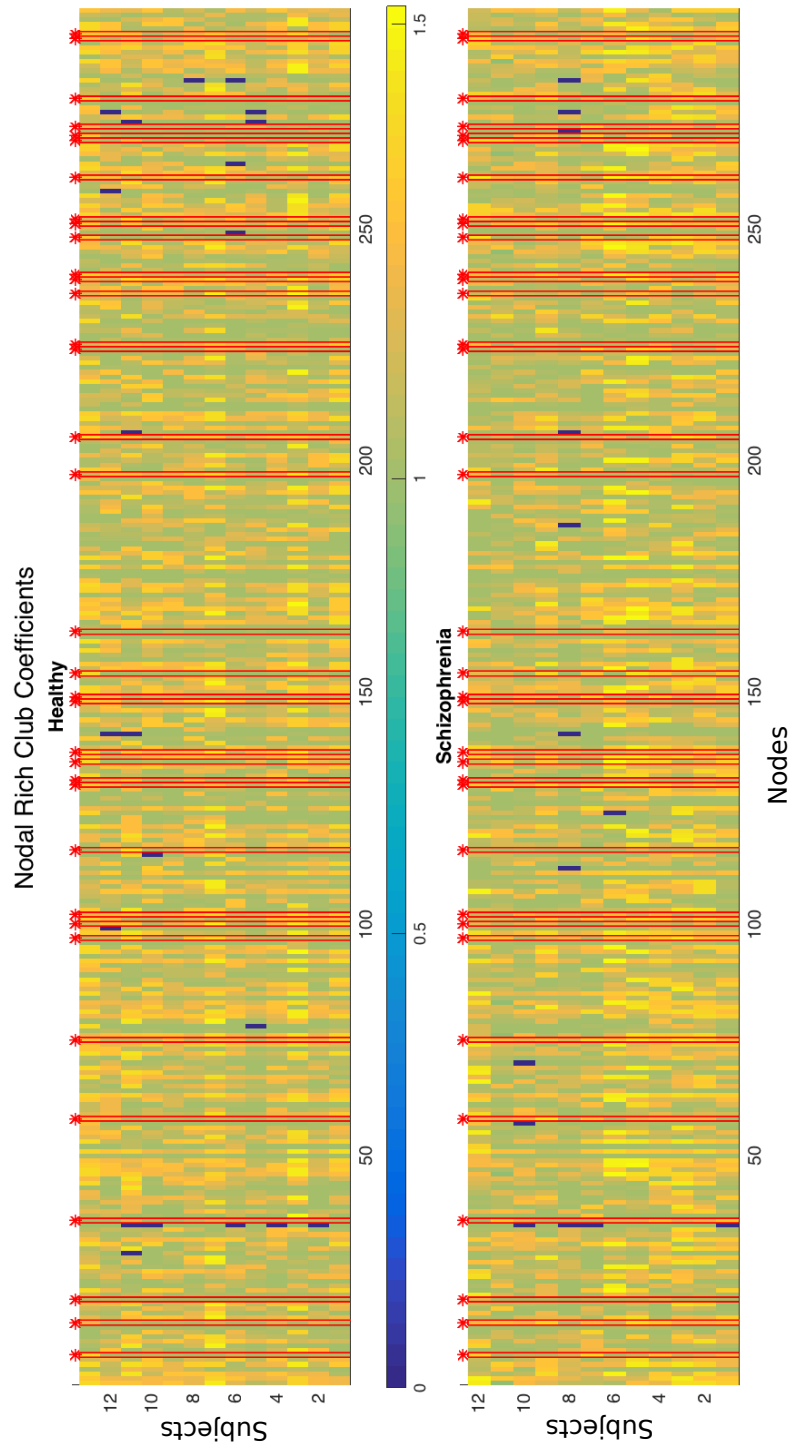


Figure 3.14: NRCC of healthy (Top) and schizophrenia (Bottom), as described in section 3.1.9. Nodes that were identified as significantly different were marked with a red asterisk

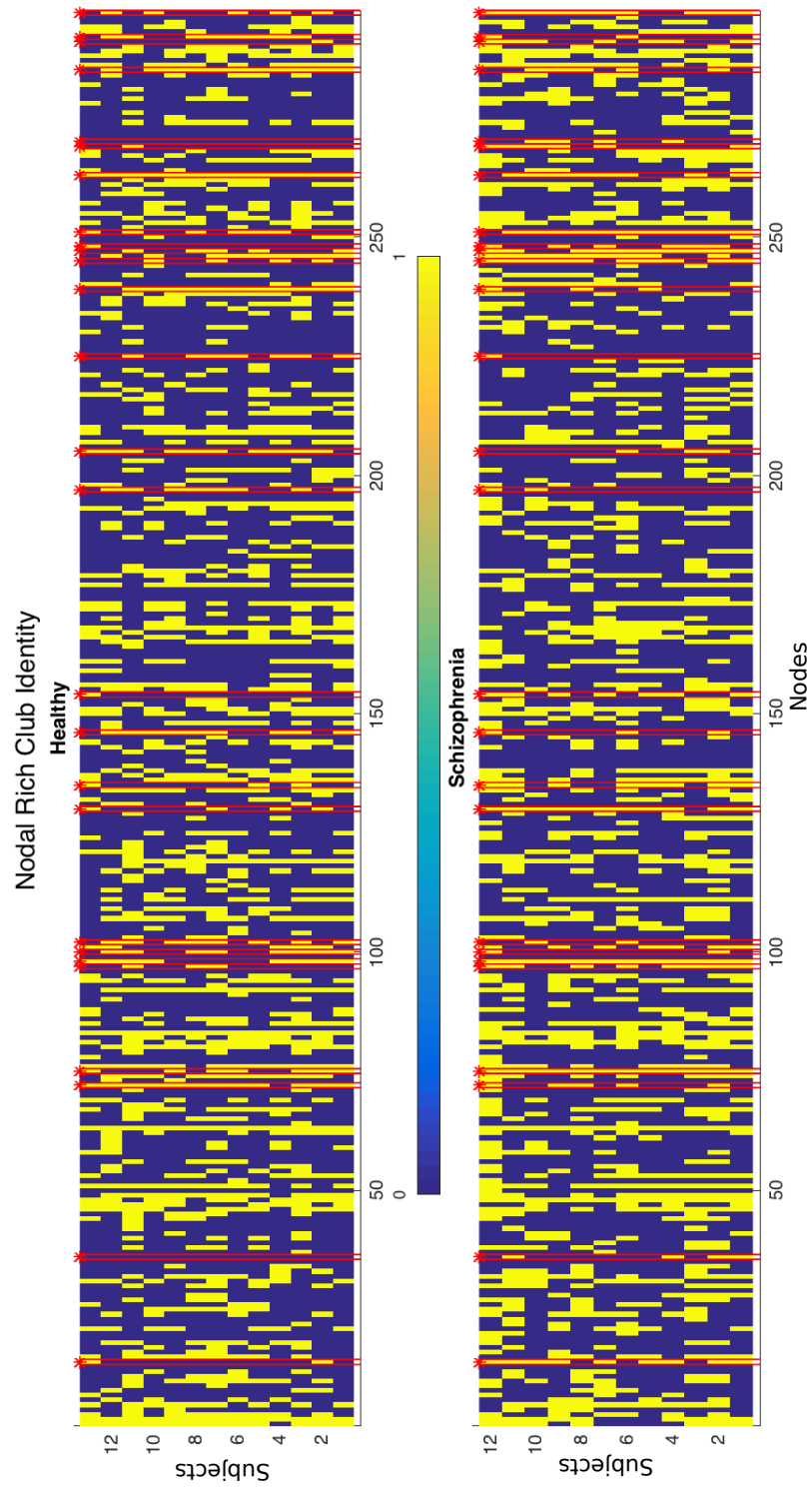


Figure 3.15: NRCI of healthy (Top) and schizophrenia (Bottom), as described in section 3.1.9. Nodes which were identified as significantly different were marked with a red asterisk

Results of the Fisher Exact test suggest that parts of the Precuneus, Insula, Cingulate and Precentral regions are no longer part of the RC organisation in schizophrenia, whereas, new regions mainly from the Cerebellum including the Declive, Culmen, Uvula, Tuber as well as a small portion of the Parietal lobe, such as parts of Inferior Parietal lobe and Angular gyrus, appear to be new members of schizophrenic RC organisation.

Additionally, we also examine how this RC organisation is consistent across the subject of each group. We use NMI (Danon *et al.*, 2005) to measure the level of similarity between the RC organisation of each pair of subjects within each group.

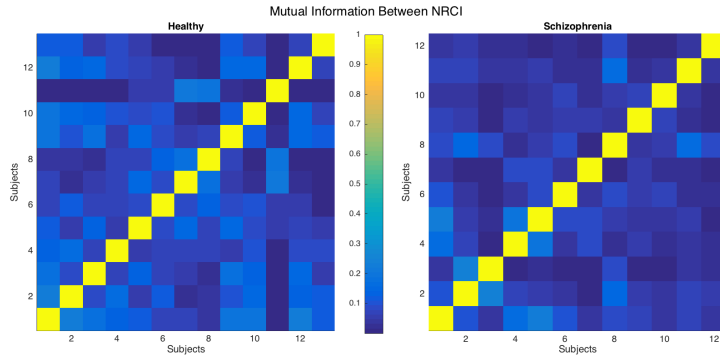


Figure 3.16: Similarity between Rich Club Identity matrices across individuals across healthy (Left) and schizophrenia (Right)

The results, in figure 3.16, suggest a relatively low consistency across subjects as the mean MNI is 0.08 and 0.05 for health and schizophrenia, respectively. However healthy subjects demonstrate a significantly higher level of consistency in comparison to schizophrenic subjects.

3.4 Discussion

In this study, we start by summarising the main findings about the topology of the functional connectome of the human brain. Furthermore, to investigate the healthy brain, we also discuss how these topological features can be used as bio-markers to identify changes during Schizophrenia.

By applying a clustering coefficient and local efficiency we showed that the human brain is more segregated than an equivalent random graph, which suggests that brain regions are in tight intra-connection with each other. In addition to the segregation measure, we also used path length and its normalised version to investigate the level of integration within the human brain. The results of characteristic path length and global efficiency suggest that the functional connectome of the human brain exhibits a level of integration as

high as its corresponding random network. This leads us to conclude that the human brain enjoys a constant balance between segregation and integration. This balance is one of the fundamental evidences for considering the human brain connectome as a complex system. The level of complexity within the human brain is measured by small-worldness which was proven to be a useful tool to summarise the topological features of a network. The global features that we discussed in this study strongly align with previous findings in the network neuroscience literature (Bullmore and Bassett, 2011; Sporns, 2011b; Sporns and Honey, 2006).

Besides the clustering coefficient and local efficiency, one additional important measure of segregation that we discuss in this study is modularity. By applying a modularity detection algorithm, we showed that the functional connectome of the human brain has a modular structure meaning that the human brain is formed by a combination of modules which are tightly intra-connected and sparsely inter-connected with each other. More importantly, results of the modularity algorithm on functional data suggest that without any prior knowledge of brain structure, the modules are spatially consistent. Our findings about the modular structure of the human brain and close association between its function and structure are widely similar to other studies in the field (Meunier *et al.*, 2010, 2014).

Further examination of the functional connectome in this study reveal that the key to emergence of complexity in the functional connectome is the collection of degree hubs which facilitates the integration among the modules. Throughout this study, we use degree centrality and betweenness centrality to detect these 'elite' nodes. Nodes among the traditional the Default Mode Network such as the majority of regions from the bilateral Precuneus, Cingulate and bilateral Inferior Parietal Lobes. These findings widely overlap with other studies of hubs in the functional connectome (Crossley *et al.*, 2014; Sporns *et al.*, 2007). In addition to identifying the degree hubs, we also examine the interactions between these influential nodes. We use the Rich Club to measure the level of intra-connection between degree hubs. By investigating the Rich Club organisation across the functional connectome we conclude that the Rich Club organisation plays a vital role in efficiency of the human brain. It is likely that this important role is due to the spatial configuration of the RC nodes. We showed that the RC nodes are geographically spread over the brain. These findings are consistent with other studies which suggest that the RC organisation plays role of a backbone to the human brain (van den Heuvel *et al.*, 2012b; Grayson *et al.*, 2014).

Although we showed how graph theoretical measures can be used to describe macro-scale human functional connectome, it is also essential to explore how these measures can be used as bio-markers for detecting changes during Schizophrenia. Prior to investigating the changes in topological features, we examined how the edge weights of un-thresholded networks change during the disease. Understanding the changes in edge weights can be

helpful to interpret the changes which are later identified in topological features. The results suggest that connections in the Parietal, Insular and Frontal lobes experience a decrease during schizophrenia. Decrease in the strength of connections in these regions were already reported in studies with different cohorts and pre-processing steps (Zalesky *et al.*, 2010, 2012b). A decrease in connectivity strength may result in the connection to be able to survive the density thresholding and therefore cause a reduction in the degree of relevant regions.

We continue examining the changes during schizophrenia by comparing the integration levels. Results of integration measures suggest that there are no differences between healthy and schizophrenic brain networks. Further to the integration level, we also investigated the changes in segregation level between healthy and schizophrenic functional connectome. Examining the Clustering coefficient and Local Efficiency suggests a decrease in the level of segregation during schizophrenia. The decrease in the level of segregation aligns with other studies in the field (Liu *et al.*, 2008; Alexander-Bloch *et al.*, 2010). However, it is notable that the normalised clustering coefficient, Ω , suggests the opposite. This contradiction may suggest that schizophrenic brain networks are more distant from their random networks which make the schizophrenic brains close to a lattice network. Eventually, we summarise the normalised integration and segregations measures by the measure of small-worldness. Our findings suggest that the small-worldness is higher in the functional connectome of schizophrenic brains.

Comparing the modular structures between group-averaged healthy and schizophrenic brain networks also suggests a larger number of clusters in schizophrenia, however, the individual modularity indexes suggest no differences between the two groups. The majority of nodes which were found to be members of the Occipital and Frontal lobes, module 3 and module 4, are represented in module B and module D, respectively. It is notable that module B in schizophrenia also accommodates other nodes of Temporal lobes. The majority of nodes in the remaining modules of healthy subjects, module 1 and module 2, share their nodes with every other module in schizophrenia.

Besides investigating the changes in global architectural features of the functional connectome during schizophrenia, we also used disruption plots to examine how the centrality measures as well as measure of local segregation changes during schizophrenia. Results of the degree centrality suggest that degree hubs lose their prominent role. These regions are mainly parts of the conjunction of the Default Mode Network and Parietal lobe including the Precuneus and parts of the Cingulate cortex. These results may routes back to the reduction in edge weights that explained earlier. In contrast, regions of the Cerebellum which were considered as peripheral nodes in healthy subjects gained more degrees during schizophrenia. Our findings of abnormalities in the role of the degree hubs is consistent

with other studies across modalities (Rubinov and Bullmore, 2013; Collin *et al.*, 2013; Lynall *et al.*, 2010). This pattern of changes can also be seen in local segregation measures where the Cuneus and Postcentral lose their local efficiency and the Cerebral Tonsil experiences an increase in its Local Efficiency. Finally, the last local measure of centrality, betweenness, suggests that nodes from the Default Mode and Saliency functional networks, such as Insular and Precuneus and Thalamus, are subject to a decrease in their Betweenness centrality, while, The Anterior Cingulate, bilateral Superior Frontal regions experience an increase.

Eventually, we propose a method to conduct multi-subject group inference between the Rich Club organisation and the coefficients of two groups. In terms of the Rich Club coefficients, the nodes that mostly experienced changes during schizophrenia are among nodes of the Parietal lobe, Frontal lobe and Cerebellum. The Rich Club coefficients in the first two lobes, Parietal and Frontal, experienced a significant decrease and the latter cortex, the Cerebellum, experienced a significant increase in its Rich Club coefficients. Similarly, the examining the Rich Club organisation also reveals that the nodes which mainly overlap with the Default Model Network, including the nodes of the Frontal and Parietal lobes, are no longer members of the Rich Club organisation in schizophrenia, instead, member of the Cerebellum cortex joins the Rich Club organisation in schizophrenic brain networks.

As it is essential to investigate the role of the Rich Club nodes in integration and centrality of brain networks, we also found that the Rich Club nodes have higher density in comparison to local and feeder regions of the networks. Betweenness centrality of the Rich Club nodes is also significantly larger across both groups. Notably, local efficiency of the Rich Club nodes in healthy subjects is significantly higher than the corresponding nodes in schizophrenia. These results suggest that the Rich Club organisations in schizophrenia play a relatively lower role in efficiency of the human brain during schizophrenia. It is notable that similar results regarding alteration of RC organisation, but in regards to structural connectivity, were reported by (van den Heuvel *et al.*, 2013a; Collin *et al.*, 2014a).

3.4.1 Summary

In this chapter, we discussed empirical graph theoretical measures which can be used to describe the macroscale connectome of the human brain in healthy and diseased conditions.

- We used total of 25 subjects, 13 healthy subjects and 12 schizophrenic subjects. A range of graph theoretical metrics were applied on the both subject and group-average levels.
- The Parietal lobe were identified as the lobe which experienced the highest level of disruption, in terms of connection strength, during schizophrenia.

- In terms of the global topological features, schizophrenic brains demonstrate a higher level of segregation which results in a higher level of small-worldness (1.43 and 1.6 in healthy and schizophrenic subjects, respectively).
- Results for another measure of segregation, modularity, suggests that the level of modularity remains the same between the two group, however the module arrangements changes due to the disorder (For instance, healthy subjects demonstrate four modules, whereas, the schizophrenic group demonstrates six modules.).
- Results for the local measures, in particular degree centrality, suggest that during schizophrenia the degree hubs lose their prominent roles and, instead, local nodes experience a rise in number of their connections (For instance, on subject level, the disruption index is -0.27).
- Two statistical frameworks were proposed to investigate the changes in organisation and coefficient, NRCI and NRCC respectively, during the disorder.
- Using the proposed frameworks, the results show that the regions, mainly related to the Parietal lobe and the DMN, experience a decrease in their RC coefficients, whereas, some regions of the Cerebellum gain some strength in their RC coefficients.

Chapter 4

Understanding a Rich Club with a Stochastic Block Model: A Connectome Study of Resting-State Schizophrenia

Previously, in section 3.1.5, we discussed a modularity algorithm which attempts to group the nodes that are more intra-connected than inter-connected. However, it is naive to assume that networks are formed merely by modular structures. Networks were shown to have more complicated patterns of connectivity which require more sophisticated clustering methods in order to be discovered. A *Stochastic Block Model (SBM)* is one of the clustering methods which subdivides network into smaller groups of nodes that exhibit homogeneous behaviour. Therefore, modular structures, can be seen as one special case of network structure which can be detected by SBMs. Since the hidden structures within a network are normally unknown, SBM can also be considered as a unsupervised learning technique.

Broadly speaking ¹, SBM techniques aim to model networks by either a single graph or collection of additive random graphs. Early attempts at using random graphs to model networks date back to Erdős-Rényi model (Erdős and Rényi, 1959), which assumes that nodes have an equal and independent probability of forming edges with other nodes. Degree distribution of this model can be described by a Binomial distribution and, therefore, can be approximated by a Poisson distribution in a network of large size but sparse connections. Although the ER model is still the most studied form of random graph, it was

¹We discussed the course of evolution of random graphs in section 2.2, however, we believe it is crucial to reiterate through a brief background of random graphs in context of SBM in the beginning of this chapter.

shown that real-life networks are of a far more complex structure to be merely modelled via a single Poisson distribution (Newman, 2010). Alternatively, scale-free networks (Barabási *et al.*, 2009) assume that majority of nodes within a network have a low number of connections while a few nodes demonstrate a large number of connections. Respectively, in the BA growth model, new nodes tend to be connected to hubs rather than peripheral nodes. Despite promising results of BA models to describe natural phenomena, this model is still far from ideal in modelling real-life complex systems (Daudin *et al.*, 2007; Stumpf *et al.*, 2005). For instance, degree distribution of the human brain was shown to follow a semi scale-free (Humphries *et al.*, 2006; Bullmore and Sporns, 2012b; Sporns and Zwi, 2004), therefore, they can neither be modelled by an ER model, since ER model fails to account for the structures, nor can they be perfectly be fitted to a scale-free model since number of hubs and their intensity are not as many as to be fitted to a power-law distribution since, as discussed in section 2.2, networks that follow power-law distribution demonstrate a very few number of highly connected nodes and many other nodes with a few connections. This motivates the need to use a newly proposed class of SBMs that subdivides a network to a collection of blocks.² In the context of SBMs, degree distribution of each block follows a Poisson distribution where the Poisson parameter varies depending on intra- and inter-connection rate of each block. This class of SBMs, conventionally known as mixture models (McLachlan and Peel, 2004), provide a unique opportunity for the modelling of a wide range of network behaviours due to its flexibility in estimating the empirical degree distribution (Mariadassou *et al.*, 2010). For instance, it was shown that SBMs are able to deliver promising results in modelling the degree distribution of C.Elegans³ (Pavlovic *et al.*, 2014).

In the context of SBMs, the degree centrality of each node has a close association with their pattern of connectivity. For instance, two nodes with a high rate of connectivity are more likely to be connected to similar nodes. Therefore, it is more likely that nodes with high degree centrality are grouped together. This realisation was the primary motivation for the author of this thesis to propose the *Rich Block (RB)*, which is a combination of SBMs with a conventional graph theoretical measure, Rich Club⁴. Similar to the Rich Club, the Rich Block aims to detect the presence of highly intra-connected network cores, which have not occurred simply due to chance.

In this study, among several forms of SBMs, we use a newly proposed complementary extension to the traditional Erdős-Rényi Mixture Model which was designed to summarise multi-subject variability between block settings of subjects in a group. The ability to account for within-group variability provides Rich Block with a unique opportunity

²Throughout this chapter we use blocks for subdivision of a network. The term 'block' was chosen to differentiate between module which were discussed in section 3.1.4.

³See section 2.3.1 for more details about topological features of the C.Elegans micro-scale connectome.

⁴To see a detailed discussion about Rich Clubs see section 3.1.9

to be seen as the first tool to tackle a long-standing challenge of multi-subject inference between rich cores of a network. Additionally, we also show how SBMs can be exploited to deliver a deeper insight into an understanding of degree hubs within a network by proposing a method, called *Degree Exceptionality (DE)*. The analysis in this chapter are carried out on both subject level, however, for sake of the visualisation we also present some group-level analysis.

4.1 Methods

In this section, we first review the basic components of the SBM model along with the estimation of its parameters and the estimation of a total number of blocks. We then briefly review some properties of the SBM model which allows us to estimate the observed network degree distribution. The approximation of the observed degree distribution is based on the cluster structure and parameters obtained from the model and in a broader sense it can be viewed as an additional goodness of fit tool to validate of the fitted clusters. Next, we briefly describe a multi-subject extension of this model, the Heterogeneous SBM (Het-SBM), which allows one to account for heterogeneous covariate effects on the subject-specific networks. Additionally, after a brief description about a Rich Block, we show how Rich Block coefficients and organisations can be found using a Het-SBM. Finally, we describe how Degree Exceptionality can be used to detect the degree hubs of a network.

4.1.1 Stochastic Block Model (SBM)

In the Stochastic Block Model (SBM) a single undirected edge between the nodes V_i and V_j is denoted by a random variable X_{ij} and, thus, an undirected network with n nodes is denoted by an $n \times n$ random and symmetric matrix $\mathbf{X} = ((X_{ij}))_{1 \leq i \neq j \leq n}$. In such a network, the set of nodes V_1, \dots, V_n is divided into Q non-overlapping blocks and this block assignment is specified by a binary random variable Z_{iq} which takes the value 1 if the node V_i is in the block q and 0 otherwise. For all of the nodes and Q the number of blocks, the $n \times Q$ random matrix \mathbf{Z} is assumed to contain independently distributed random variables \mathbf{Z}_i each of which follows a Categorical distribution with Q possible outcomes.

$$\mathbf{Z}_i \sim \text{Categorical}(Q, \boldsymbol{\alpha}), \quad (4.1)$$

where $\boldsymbol{\alpha}$ is a $1 \times Q$ vector of success probabilities such that $\sum_{q=1}^Q \alpha_q = 1$. Conditional on the nodes' block assignments, the edges are assumed to follow a Bernoulli distribution

$$X_{ij}|Z_{iq} = 1, Z_{jl} = 1 \sim \text{Bernoulli}(\pi_{ql}) \quad (4.2)$$

where a single parameter π_{ql} gives probability of an edge between blocks q and l . Thus, the connectivity rates for all the blocks in a network are given by a $Q \times Q$ matrix $\pi = ((\pi_{ql}))_{1 \leq q, l \leq Q}$. It is interesting to point out that since block edges are homogeneously distributed within that block, the clustered network can be also seen as $\binom{Q+1}{2}$ number of classic Erdős-Rényi models. This is why in the literature this classical SBM model is denoted by the Erdős-Rényi Mixture Model (ERMM) (Zanghi *et al.*, 2008). To summarise, it is crucial to reiterate that the model is constructed from α and Π and estimating these parameters are dependent on another latent parameter, Z . In the following section, we explain how to estimate the latent parameter Z and, consequently, the model parameters α and Π .

Estimation and model selection

In the SBM, the likelihood of the complete data (i.e. (X, Z)) can be stated as

$$\log f(\mathbf{x}, \mathbf{z}; \pi, \alpha) = \log f(\mathbf{x}|\mathbf{z}; \pi) + \log f(\mathbf{z}; \alpha), \quad (4.3)$$

and because \mathbf{z} also needs to be estimated, the optimisation is carried out on the marginal likelihood of the complete likelihood ($\log f(\mathbf{x}; \pi, \alpha)$). However, as the computation of the marginal likelihood involves a search over all possible realisations of \mathbf{Z} this quickly becomes computationally challenging which is why a variational approximation is used instead (Nowicki and Snijders, 2007; Daudin *et al.*, 2004). In the variational approach (Jordan *et al.*, 1999), the model parameter estimates are found from the variational bound $\mathcal{J}(f^*(\mathbf{z}; \tau); \pi, \alpha)$ defined as

$$\mathcal{J}(f^*(\mathbf{z}; \tau); \pi, \alpha) = \mathbb{E}_{f^*}[\log f(\mathbf{x}, \mathbf{Z}; \pi, \alpha)] - \mathbb{E}_{f^*}[\log f^*(\mathbf{Z}; \tau)], \quad (4.4)$$

where the density $f^*(\mathbf{z}; \tau)$ depends on the variational parameter τ and it minimises the Kullback-Leibler distance to $f(\mathbf{z}|\mathbf{x}; \pi, \alpha)$ (see Appendix B.1). The natural choice for $f^*(\mathbf{z}; \tau)$ is a Categorical distribution with independent block success probabilities for each node, such that

$$f^*(\mathbf{z}; \tau) = \prod_{i=1}^n \prod_{q=1}^Q \tau_{iq}^{z_{iq}}, \quad (4.5)$$

where $\sum_{q=1}^Q \tau_{iq} = 1$ and $\mathbb{E}_{f^*}(Z_{iq}) = \tau_{iq}$ and $\mathbb{E}_{f^*}(Z_{iq}Z_{jl}) = \tau_{iq}\tau_{jl}$. Although the computation of the variational bound $\mathcal{J}(f^*(\mathbf{z}; \tau); \pi, \alpha)$ (see Eq. 4.4) involves expectations with respect to \mathbf{Z} , these quantities are generally straightforward to obtain as the \mathbf{Z} 's are mutually independent. Maximising Eq. (4.4) with respect to τ , α and π yields the following point estimate

on equations:

$$\hat{\tau}_{iq} \propto \hat{\alpha}_q \prod_{j \neq i}^n \prod_{l=1}^Q [\hat{\pi}_{ql}^{x_{ij}} (1 - \hat{\pi}_{ql})^{1-x_{ij}}]^{\hat{\tau}_{jl}}, \quad (4.6)$$

$$\hat{\alpha}_q = \frac{1}{n} \sum_{i=1}^n \hat{\tau}_{iq}, \quad (4.7)$$

$$\hat{\pi}_{ql} = \frac{\sum_{i \neq j}^n \hat{\tau}_{iq} \hat{\tau}_{jl} x_{ij}}{\sum_{i \neq j}^n \hat{\tau}_{iq} \hat{\tau}_{jl}}. \quad (4.8)$$

In this model, the optimal number of blocks Q is estimated via the Integrated Classification Criterion (ICL) (Biernacki *et al.*, 2000), which is applied to the completed data likelihood (see Eq. 4.3) (Daudin *et al.*, 2004, 2007). In particular, the estimate of Q is obtained by comparing the ICL scores between the models with a different / the same number of blocks and different parameter estimates. For a model \mathcal{M}_Q with Q number of blocks, the ICL criterion is given as

$$\text{ICL}(\mathcal{M}_Q) = \underbrace{\arg \max_{(\alpha, \pi)} \log f(\mathbf{x}, \hat{\mathbf{z}} | \mathcal{M}_Q, \hat{\alpha}, \hat{\pi})}_{\text{complete data log likelihood}} - \underbrace{\frac{1}{2} \frac{Q(Q-1)}{2} \log \left[\frac{n(n-1)}{2} \right] - \frac{Q-1}{2} \log[n]}_{\text{trade-off term}}, \quad (4.9)$$

where the first under-braced term is the complete data log likelihood evaluated at the variational estimates of $\hat{\alpha}$ and $\hat{\pi}$, where the latent \mathbf{z} is replaced by its estimate, $\hat{\mathbf{z}}$. The second under-braced term balances the trade-off between the goodness of fit and the complexity of the model. In particular for large values of Q , this term becomes large and consequently it penalises the complete data log likelihood more severely. In this way, the parsimony of the selected model is preserved.

Degree Distribution

The degree $\rho(V_i)$ of a node V_i is defined as the total number of edges which terminate at V_i . Formally, this can be written as $\rho(V_i) = \sum_{j \neq i}^n X_{ij}$. In particular, $\rho(V_i)$ is a random variable while ρ_i is its realisation. Given the block assignment \mathbf{Z} and given that the vertex V_i is a member of block q , the conditional distribution of the degree of node V_i can be written as:

$$\rho(V_i) | Z_{iq} = 1 \sim \text{Binomial}(n-1, \bar{\pi}_q) \approx \text{Poisson}(\lambda_q), \quad (4.10)$$

$$\bar{\pi}_q = \sum_{\ell=1}^Q \hat{\alpha}_\ell \hat{\pi}_{q\ell} \quad \lambda_q = (n-1) \bar{\pi}_q, \quad (4.11)$$

where $\bar{\pi}_q$ is a weighted sum of connection rates between block q and the rest of the network and λ_q is the expected degree of block q (see Appendix B.2 for more details). This implies

that the probability of a particular degree value ρ_i of a node V_i can be expressed as a mixture of Poisson (or Binomial) densities depending on our choice of $P(\rho_i|Z_{iq} = 1)$ (see Eq. 4.11), since:

$$\begin{aligned} f(\rho_i) &= \sum_{l=1}^Q P(\rho_i|Z_{iq} = 1)P(Z_{iq} = 1), \\ &= \sum_{l=1}^Q \hat{\alpha}_q P(\rho_i|Z_{iq} = 1). \end{aligned} \quad (4.12)$$

Density of the Fitted ERMM

The SBM parameters allow us to obtain the expected edge count between different blocks. Given a block assignment \mathbf{Z} and by defining the number of edges between block q and l as $E_{ql} = \frac{1}{2} \sum_{i=1}^n \sum_{j \neq i} Z_{iq} Z_{jl} X_{ij}$, we can state its expectation as:

$$\mathbb{E}(E_{ql}) = \frac{n(n-1)\hat{\alpha}_q \hat{\alpha}_l \hat{\pi}_{ql}}{2}, \quad (4.13)$$

(see Appendix B.2 for more details). Similarly, noting that

$$f(x_{ij}) \propto \sum_{q=1}^Q \sum_{l=1}^Q \hat{\alpha}_q \hat{\alpha}_l P(X_{ij} = x_{ij}|Z_{iq} = \hat{z}_{iq}, Z_{jl} = \hat{z}_{jl}), \quad (4.14)$$

we can obtain the expectation of the density $\mathbb{E}(X_{ij})$ as

$$\begin{aligned} \mathbb{E}(X_{ij}) &= \sum_{i=1}^n \sum_{j < i}^n x_{ij} \sum_{q=1}^Q \sum_{l=1}^Q \hat{\alpha}_q \hat{\alpha}_l P(X_{ij} = x_{ij}|Z_{iq} = 1, Z_{jl} = 1) \\ &= \sum_{q=1}^Q \sum_{l=1}^Q \hat{\alpha}_q \hat{\alpha}_l \sum_{i=1}^n \sum_{j < i}^n x_{ij} P(X_{ij} = x_{ij}|Z_{iq} = 1, Z_{jl} = 1) \\ &= \sum_{q=1}^Q \sum_{l=1}^Q \hat{\alpha}_q \hat{\alpha}_l \hat{\pi}_{ql}. \end{aligned} \quad (4.15)$$

4.1.2 Generalised Linear Stochastic Block Model (GL-SBM)

The ERMM was shown to be promising in the study of single subject analysis (Daudin *et al.*, 2004). However, in the analysis of multi-subject networks, the traditional ERMM estimates a different block assignment for each subject. Varying block assignment across subjects makes the inference between subjects impossible. Three extensions to the ERMM, which can be useful in multi-subject studies of block models, have recently been proposed by Pavlovic *et al.* (2015). These extensions are from a family of Generalised Linear Stochastic

Block Models (GL-SBM), which aim to estimate the most common block assignments between all subjects. The first and most straightforward of these three models is the Binomial Stochastic Block Model (Bin-SBM) which assumes that there is no subject-specific variability between subjects. The second method, the Homogeneous Stochastic Block Model (Hom-SBM), assumes that variabilities between subjects have a global effect on each subject, and therefore, affects overall density of the network by either increasing or decreasing homogeneously. The last and the most sophisticated method is the Heterogeneous Stochastic Block Model (Het-SBM) which assumes that variabilities between subjects affect regions (inter-block and intra-block connectivities) of the connectivity matrix differently. As we are interested in using a subject-independent threshold and, in the meantime, we are also interested in the effect of variabilities between subjects, we only use a Het-SBM in this work.

In the Het-SBM, there is a total of K subjects and each subject is associated with a single network. Thus, for the k -th subject the edge between the nodes V_i and V_j is defined as X_{ijk} . Conditional on the node assignments, the edges follow a Bernoulli distribution such that each subject's block parameters π_{qlk} depend on a logistic regression model as:

$$X_{ijk}|Z_{iq} = 1, Z_{jl} = 1 \sim \text{Bernoulli}(\pi_{qlk}) \quad (4.16)$$

$$\log\left(\frac{\pi_{qlk}}{1 - \pi_{qlk}}\right) = \mathbf{d}_k^T \boldsymbol{\beta}_{ql} \quad (4.17)$$

where \mathbf{d}_k^T is a $1 \times P$ dimensional vector of covariates related to the k -th subject and $\boldsymbol{\beta}_{ql}$ is a vector of size $1 \times P$ of regression parameters. For a fixed Q , the variational bound can be written as

$$\begin{aligned} \mathcal{J}(f^*(\mathbf{z}; \boldsymbol{\tau}); \boldsymbol{\alpha}, \boldsymbol{\beta}) = & \frac{1}{2} \sum_{k=1}^K \sum_{i=1}^n \sum_{j \neq i}^n \sum_{q \leq l}^Q \gamma_{ijql} \log[f(x_{ijk}|z_{iq}, z_{jl}; \boldsymbol{\beta}_{ql})] \\ & + \sum_{i=1}^n \sum_{q=1}^Q \tau_{iq} \log[\alpha_q] - \sum_{i=1}^n \sum_{q=1}^Q \tau_{iq} \log[\tau_{iq}] \end{aligned} \quad (4.18)$$

where γ_{ijql} , is defined as

$$\gamma_{ijql} = \begin{cases} \tau_{iq} \tau_{jq} & \text{if } q = l \\ \tau_{iq} \tau_{jl} + \tau_{il} \tau_{jq} & \text{if } q < l. \end{cases} \quad (4.19)$$

We note the equivalence between $\sum_{i=1}^n \sum_{j \neq i}^n \sum_{q,l}^Q \tau_{iq} \tau_{jl}$ and $\sum_{i=1}^n \sum_{j \neq i}^n \sum_{q \leq l}^Q \gamma_{ijql}$. An optimi-

sation of this bound leads to the following point estimating equations:

$$\hat{\tau}_{iq} \propto \hat{\alpha}_q \prod_{k=1}^K \prod_{j \neq i}^n \prod_{l=1}^Q \left[\frac{e^{x_{ij} d_k^T \hat{\beta}_{ql}}}{1 + e^{d_k^T \hat{\beta}_{ql}}} \right]^{\hat{\tau}_{jl}} \quad (4.20)$$

$$\hat{\alpha}_q = \frac{1}{n} \sum_{i=1}^n \hat{\tau}_{iq}. \quad (4.21)$$

To account for the effect of small samples we use Firth type estimates (Firth, 1993), which are based on the regularised variational bound

$$\mathcal{J}^*(f^*(z; \tau); \alpha, \beta) = \mathcal{J}(f^*(z; \tau); \alpha, \beta) + \frac{1}{2} \sum_{q \leq l}^Q \log[\text{Det}(\mathcal{I}_{ql}(\beta_{ql}))], \quad (4.22)$$

where $\mathcal{I}_{ql}(\beta_{ql})$ is a Fisher information matrix for block (q, l) obtained from Eq. 4.18. The estimates of β_{ql} are found by the Newton-Raphson method so that for the m -th iteration it can be defined as:

$$\beta_{ql}^{(m)} = \beta_{ql}^{(m-1)} + \mathcal{I}_{ql}^{-1}(\beta_{ql}^{(m-1)}) \mathcal{U}_{ql}^*(\beta_{ql}^{(m-1)}), \quad (4.23)$$

where \mathcal{U}_{ql}^* is a score function associated with the block (q, l) obtained from Eq. 4.22.

Eventually, to select the best model, we use the ICL criterion described in Section 4.1.1. The ICL for this model can be stated as

$$\text{ICL}(\mathcal{M}_Q) = \arg \max_{(\alpha, \beta)} \log f(\mathbf{x}, \mathbf{z} | \mathcal{M}_Q, \hat{\alpha}, \hat{\beta}) - \frac{1}{2} \left(\frac{Q(Q+1)}{2} P \right) \log \left(\frac{n(n-1)}{2} K \right) - \frac{Q-1}{2} \log[n]. \quad (4.24)$$

4.1.3 Rich Block

A Rich Block aims to describe the interaction between highly connected blocks of a network. As shown in Eq. 4.11, within the context of a SBM, the expected degree of a block is determined by a mixture of inter- and intra- connectivity rates of the block and its size. A Rich Block organisation of expected degree k is formed by the connection among collection of mini-blocks⁵ with expected degree k or larger; therefore, the connection matrix of this configuration is defined as $\mathbf{\Pi}_{P \times P}^{RB}$, where P is cardinality of set of mini-blocks of blocks with expected degree of k or larger. In this setting, the number of RB organisation is equal to number blocks which were found by SBMs. It is worth noting the first RB coefficient is intra-connection rate of the block with the highest expected degree and the last RB

⁵Each block is formed of collection of mini-block which determines the rate of connectivity between the block and other blocks. Inter-mini-blocks sits at the intersection between the block and the diagonal of the connectivity matrix which represent the connectivity rate of nodes within that block. Intra-mini blocks are off diagonal blocks representing the connectivity between two septate blocks.

coefficient is the expected density of the network.

As discussed in section 3.1.9, Rich Clubs were shown to be sensitive to variations in the degree centrality of nodes of a subject, and therefore, it is rarely the case that the Rich Club organisation is consistent across subjects. Consequently, this makes group inference between RC of subjects impossible, as the changes in degree centrality of nodes across subjects can change the club membership of nodes. However, by using the parameter Z , obtained from Het-BMM, as the Rich Block assignment we can rectify this problem since the Het-SBM estimates a common block assignment across subjects therefore the difficulties in group inference between network cores merely boils down to comparing the connection rates of the Rich Blocks.

It is notable that blocks with higher rate of inter-connectivity are more likely to form additional connections, purely based on chance, with other highly inter-connected blocks, which makes it troublesome to examine genuine RB effects. Therefore, the RB coefficients should be compared by an appropriate null model. This leads us to divide the analysis of the Rich Block effect into two main components:

1. *Rich Block Detection* which compares the mixture degree distribution of mini-blocks involved in a Rich Block organisation against mixture degree distribution of corresponding blocks in a null model. On each λ level if the mixture degree distribution is significantly different, then the relative RB setting is identified as a RB organisation.
2. *Rich Block Coefficients* are only concerned with estimating the rate of connectivity between blocks with expected degree, λ , or larger regardless of whether the connectivity rate is sourced in chance or the genuine RB effect.

In the following, we start by introducing an appropriate null model for Rich Block detection and then we discuss how Rich Block Organisation and Rich Block Coefficients can be identified. Fig 4.3 also illustrates the process which should be taken to estimate the organistaions and coefficients.

Null Model for Rich Blocks

In context of SBMs, a null model is defined as a connectivity rate matrix, Π_{Null} , which lacks any form of structure but where each block has an identical size and expected degree as in an estimated connectivity rate matrix, Π . It is worth noting that the null model that was defined here is quite similar to what was discussed in section 3.1.5. One intuitive way of forming null models, is to generate a large number of randomised networks using the method described in section 3.1.5 and fit SBM techniques to each randomised network. Execution time for the process explained is more than 1.20 hours (given each of SBM fitting

takes 1.1s) per a subject of 300 nodes⁶. However, this technique is not only computationally inefficient, but also the estimated block assignments between each randomised network may turn out to be different. Different block assignments in randomised networks make further statistical inferences practically impossible.

Inspired by the null model proposed by Newman (2013) to detect the modular structures in empirical networks, we propose a method, called a Tensor Null Model (TNM), to generate a randomised connectivity matrix, $\mathbf{\Pi}_{Null}$. In a TNM, the expected degree of each block is preserved while the structure within the connectivity rate matrix is eliminated by tensor multiplication of expected degrees. Assuming equal α s, null probability connectivity rate matrix is defined as:

$$\mathbf{\Pi}_{Null} = \frac{\mathbf{S} \otimes \mathbf{S}^T}{\sum \mathbf{S}} \quad (4.25)$$

where \mathbf{S} is a column vector representing the block strength of connectivity matrix, $\mathbf{\Pi}$. In the case of models with unequal block sizes, we use a weighted sum of block strengths to obtain the vector \mathbf{S} . It is worth mentioning that the TNM execution time is 0.019s which is 63k times faster than the conventional randomisation process.

Rich Block Detection

To examine whether a Rich Block organisation suggests a significantly larger rate of connectivity compared to its null model, we compare the degree distribution of Rich Block organisation with its corresponding degree distribution in the null model. Although the comparison should be conducted between the 'density' of each organisation in the estimated and null connectivity matrix, we only use degree distributions as, according to the handshaking lemma (Euler, 1741), the average degree of a network is double the number of edges.

To form the degree distribution of a Rich Block organisation, we use a finite mixture of Binomial distributions of mini-blocks that contribute to a Rich Block. Given the Z_{iq} and λ_q from section 4.1.1 we can extract the Rich Block organisation of blocks with degree equal or larger than λ_q out of the connectivity matrix $\hat{\mathbf{\Pi}}$. The Rich Block organisation of degree k , has a known block assignment ξ_{ip}^k which indicated whether node V_i is a member of a block p while block p is a member of the Rich Block organisation. Therefore, we can estimate the probability that a node drops in block p as:

$$\alpha'_p = \frac{1}{n'} \sum_{p=1}^{n'} \xi_{ip}^k = \frac{\hat{\alpha}_p}{\sum_{p=1}^{n'} \hat{\alpha}_q} \quad (4.26)$$

⁶The execution times were calculated on a OS X El Capitan with 2.5GHz process and 16G memory.

where, n' , is number of nodes in a Rich Block organisation. Therefore, the weighted sum of connections between block p and the rest of the blocks within $\mathbf{\Pi}^{RB}$ can be defined as:

$$\bar{\pi}_p' = \sum_{\ell=1}^P \alpha_p' \pi_{p\ell}. \quad (4.27)$$

Inspired by Eq. 4.12, we can approximate the degree distribution of RB organisation of λ as a finite mixture model of Binomials. Therefore, the probability of a node, E_i , from a Rich Block organisation having degree $\rho(E_i)$ can be defined as:

$$\begin{aligned} f^k(\rho_i) &= \sum_{\ell=1}^P P(\rho_\ell | \xi_{\ell p}^k = 1) P(\xi_{\ell p}^k = 1), \\ &= \sum_{p=1}^P \alpha_p' \text{Binomial}(n' - 1, \bar{\pi}_p') \end{aligned} \quad (4.28)$$

Statistical Testing For testing whether a Rich Block organisation genuinely exhibits a Rich Block effect, we approximate the degree distribution of the organisation in the estimated connectivity rate matrix, $\hat{\mathbf{\Pi}}$ and the randomised connectivity rate matrix, $\hat{\mathbf{\Pi}}_{Null}$. We then obtain the Cumulative Density Function (CDF) of each of these two degree distributions aiming to generate synthetic degrees. Then, after generating one million, where the results reach stability, synthetic degrees out of the CDF of $f^k(\rho_i)$ and $f_{Null}^k(\rho_i)$, we obtain the p-value of null hypothesis that the degree distributions are not different. The null hypothesis is rejected when less than 5% of the degrees generated in $f^k(\rho_i)$ are larger than the degrees generated in $f_{Null}^k(\rho_i)$ and consequently the Rich Block organisation of degree k is significant. A visual example of this process can be found in Fig 4.1 where the null and estimated degree distributions for RB organisation 1 and 11 were shown.

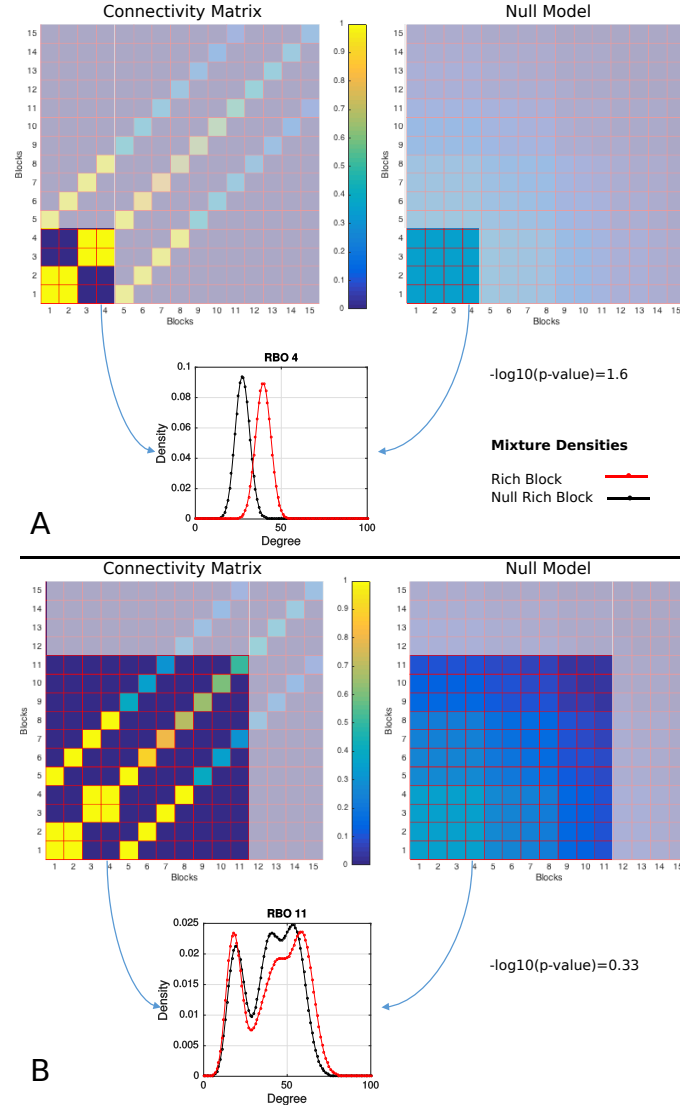


Figure 4.1: Rich Block detection procedure on a toy connectivity rate matrix, $\hat{\Pi}$, and its null model, Π_{Null} , were shown, representing a network of 300 nodes and $Q = 15$. Binomial Mixture densities used to generate degree distributions of two arbitrary RB organisation of 4 (**A**) and 11 (**B**) illustrated. The results of RB detection suggest that $H_0 : f^k(\rho_i) > f_{Null}^k(\rho_i)$ is significant, while similar hypothesis fails to be rejected on the 11th RB organisation.

Rich Block Coefficients

In section 4.1.1, we showed how the global density of a network can be obtained by SBM parameters. Here, we show how, by using Eq. 4.15, we can estimate the density of each Rich Block organisation. Given the connection rate matrix, $\hat{\Pi}$, and the prior probability, α ,

we can obtain the density of each Rich Block organisation as:

$$\Upsilon^k = \sum_{q=1}^P \sum_{\ell=1}^P \alpha'_q \alpha'_\ell \pi_{q\ell}. \quad (4.29)$$

Similar to a RC, RB coefficients should also be normalised by their respective null RB coefficients, Υ_{Null}^k , to eliminate the bias caused due to chance. Normalised Rich Block coefficient is defined as:

$$\Upsilon_{Norm}^k = \frac{\Upsilon^k}{\Upsilon_{Null}^k} = \frac{\sum_{q=1}^P \sum_{\ell=1}^P \alpha'_q \alpha'_\ell \pi_{q\ell}}{\sum_{q=1}^P \sum_{\ell=1}^P \alpha'_q \alpha'_\ell \pi_{q\ell}^{Null}}. \quad (4.30)$$

For visual example of how RB coefficients are estimated, and their comparison with RC coefficients, see Fig 4.2. From now on, the term RB coefficient is referred to a normliased RB coefficients unless, it is stated otherwise.

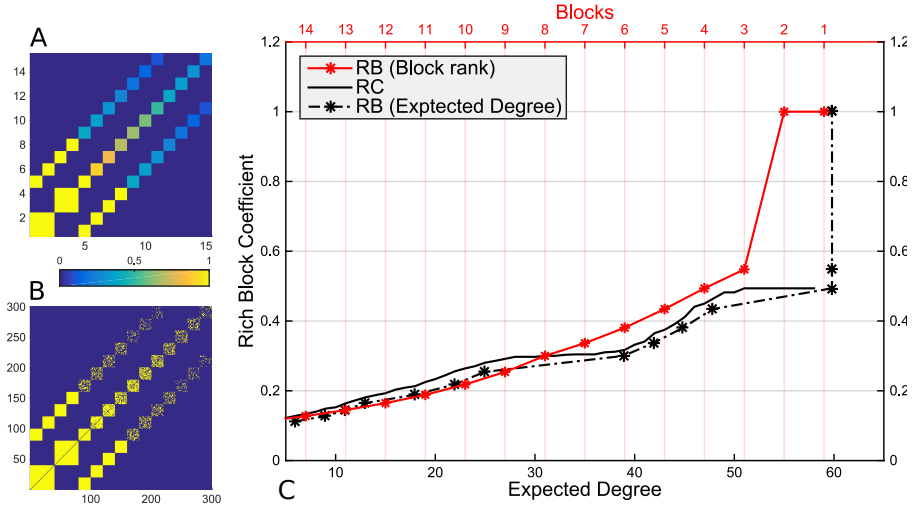


Figure 4.2: **A.** Connectivity probability matrix of a toy network with 15 (equal-sized) blocks **B.** Simulated network. **C.** shows the RB coefficients, for this network. It is notable how the RB coefficients could differentiate between the blocks with the same expected degree but different patterns of connectivity.

Feeder and Local Connections

Detecting significant RB organisation allows us to examine the interactions between the RB organisations and the rest of the network. Similar to section 3.1.9, these interactions are facilitated by a collection of mini-blocks which are classed into two main categories. The first category is *Feeder*, which represents the blocks that connect the RB organisations to

the rest of the network. In the context of RB, feeders are a collection of off-diagonal mini-blocks which participate in the expected degree of hub blocks, but they are not detected as part of a RB organisation. The second category is *Locals*, which refer to a collection of inter- or intra- block interactions that merely form between non-RB blocks. More simply, they are neither participating in the degree of a significant RB nor identified as one. We use Eq. 4.29 to calculate the connection rates in each category of interactions.

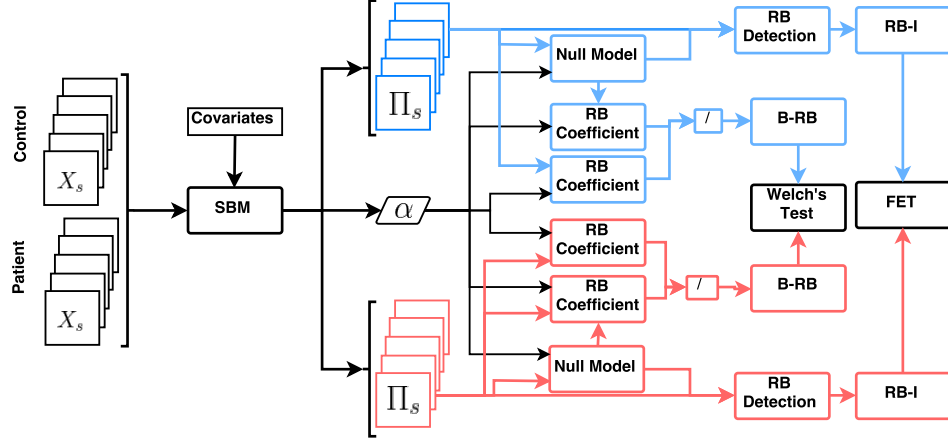


Figure 4.3: Flowchart of calculation Rich Block and group inference between two groups. Rectangles indicate process, parallelogram shows data. Blue indicate processes that should be undertaken on control and red for patient. Also note that colourful sections should be repeated for each individual, however, we only depicted this process for one Π per group.

Group Inference for Multi-subject Rich Block Coefficients and Organisation

Aiming to conduct statistical inference between RB coefficients and organisations between two groups, we propose two techniques which heavily rely on the notions of NRCC and NRCI explained in section 3.1.9:

Block-wise Rich Block Coefficient (BRBC) is a matrix of size $K \times Q$. Each element of this matrix is set with RB coefficients corresponding to the degree of the Rich Block organisations. For example, if block q in subject k has an expected degree of λ_g , then element (k,q) of BRBC matrix is set to Υ_{Norm}^g . For each group under study, we then produce two separate matrices of $BRBC^{control}$ and $BRBC^{patient}$. Finally, hypothesis testing can be applied between each column of these two matrices.

Rich Block identity (RBI) is a binary matrix of size $S \times Q$. Each element of this matrix indicates whether the corresponding block has been detected as a member of a sig-

nificant RB organisation. Therefore:

$$RBI(s, q) = \begin{cases} 1, & \text{if } P(f^k(\rho_i) \neq f^k(\rho_i)_{Null}) < 5\% \\ 0, & \text{Otherwise.} \end{cases}$$

Eventually by forming separate RBI matrices corresponding to each group, $RBI^{control}$ and $RBI^{patient}$, a hypothesis test can be applied to each column of these two matrices to identify the significant differences between the RB organisation of two groups. For this study, we use the Fisher Exact test from a class of small-sample categorical tests.

4.1.4 Degree Analysis

Although measures of centrality, such as Degree Centrality, are among some of the most widely used graph theoretical methods, finding hubs within a network is still a subject of ongoing debate due to the popular challenge of deciding on a threshold after which a node or a block is considered to be a hub. In this section, we introduce a SBM-based measure called *Degree Exceptionality* (DE) which aims to detect the hub blocks of a network. In the context of DE, a block is a degree hub when its expected degree, λ , is significantly larger than the expected degree of rest of the network. For testing whether block ℓ is a hub block, Monte-Carlo simulations were used to generate one million synthetic degrees out of the degree distribution of each block. We then tested each under the null hypothesis that the expected degree of the block λ_ℓ is significantly larger than the expected degree of a mixture of all other blocks:

$$\mathcal{H}_0 : P\left(\sum_{q \in \mathbf{B}} \alpha_q \text{Poisson}(\lambda_q) > \alpha_\ell \text{Poisson}(\lambda_\ell)\right) \quad (4.31)$$

where \mathbf{B} contains all blocks except block ℓ . Blocks with p-value lower than 5% are considered as hub blocks. Therefore, by considering Eq. 4.11, we can conclude that a block with a small number of nodes but relatively higher rate of connectivity is more likely to be detected as a DE.

For further analysis of the hub blocks, we also adopt the notion of Hub Disruption that was discussed in section 3.1.10 by substituting the former degree centrality of the nodes with expected degree of blocks. Therefore, in the context of SBM, the y-axis of the disruption plot shows the difference between the expected degree of blocks between the two groups of healthy and the patient and the x-axis shows the average of the expected values in the two groups. Thereby, when a point, corresponding to changes in a block's expected degree during a disease, appears in a negative side of the y-axis, it indicates that the rate of inter- and intra- connections experience a decrease during the disease and vice versa.

4.2 Simulation Methods

Aiming to validate the methods proposed so far, we use a synthetic connection rate matrix, Π , and a synthetic prior probability, α , to simulate networks with a different number of nodes and blocks. Given these two parameters, we generate surrogate networks according to Eq. 4.11.

For simulating the Rich Block effect within a network, regardless of network size, 25% of each connectivity rate matrix, Π , was assigned with higher intra-connection probabilities, whereas, the rest of the network is assigned with relatively lower connection rates. As it is known that the human brain network has a modular structure, diagonal blocks have a relatively larger connection rate. Since variation in the network density can drastically bias the Rich Block/Club analysis, we ensure that the generated networks have the same global density by using Eq. 4.15. In terms of the network size, every simulated network has three classes of 60, 120 and 300 nodes. Each of these classes has a sub-class with a different block number of 6, 15, 30 and 60. However, due to the low $\frac{\text{node}}{\text{block}}$ ratio (≥ 4), we later discard three of these classes (60 nodes, 30 blocks; 60 nodes, 60 block; 120 nodes, 60 block).

We also examined how successfully Degree Exceptionality can find hub blocks of networks of different nodes and block sizes. We used synthetic connection rate matrices of 10, 20 and 30 blocks. Using each of these matrices, we produced networks of size 120, 240 and 360. To simulate the effect of hubs we only set 10 % of blocks to a remarkably higher connection rate ($\approx 98\%$). Therefore, each category has 1,2 and 3 hub blocks, respectively.

Finally, we use a 15 equal-sized connection rate matrix to examine the performance of the Tensor Null Model. The connection rate matrix setting has hub blocks as well as several same-degree blocks with different pattern of connectivity.

4.2.1 Validation Methods

In this section, we describe how each of the proposed methods can be validated by simulation methods described in previous section. It is worth noting that for validation, we only use a single subject ERMM, discussed in section 4.1.1, as setting a ground-truth for the Rich Block resulting from multi-subject Het-SBM appears to be cumbersome.

Tensor Null Model

To produce a ground-truth for validating the TNM, we first use the synthetic connectivity rate matrix described in section 4.2 to generate 1k simulated networks, then using the Maslov & Sneppen (M&S) method, discussed in section 3.1.5, we randomise each of these

one thousand networks. Eventually, we use block assignments of the synthetic network to obtain a connection rate matrix for each of the randomised networks using Eq. 4.8.

Using the Kullback-Leibler divergence (Kullback and Leibler, 1951), D_{KL} , we measure the distance between the averaged degree distribution of the M&S randomisation algorithm, $f_{Null-MS}^k(\rho_i)$, and the corresponding degree distribution of the RB organisation of degree k , $f^k(\rho_i)$. We separately repeat this measurement between the degree distribution of the TNM, $f_{Null-TNM}^k(\rho_i)$ and the degree distribution for each RB organisation, using:

$$\begin{aligned} D_{KL}^k(f_{Null-MS}^k(\rho_i) || f^k(\rho_i)) &= \sum_i f_{Null-MS}^k(i) \ln\left(\frac{f_{Null-MS}^k(i)}{f^k(i)}\right) \\ D_{KL}^k(f_{Null-TNM}^k(\rho_i) || f^k(\rho_i)) &= \sum_i f_{Null-TNM}^k(i) \ln\left(\frac{f_{Null-TNM}^k(i)}{f^k(i)}\right). \end{aligned} \quad (4.32)$$

By comparing the distances between each of these randomisation algorithms, we can examine how TNM performance differs from the MS algorithm for each of the RB organisations.

Nodal Richness Coefficients

To investigate how similar Rich Block coefficients can estimate Rich Club coefficients, we generate one thousand simulated networks for each of the categories described in section 4.2. To examine the similarity between RB and RC coefficients, we construct Nodal Richness Coefficients (NRC) vectors of size $1 \times N$ for Rich Club coefficients (NRC^{RC}) and Rich Block coefficient (NRC^{RB}), separately. Each element of these two vectors has a RB/RC coefficient of the block/club that the nodes belongs to. Thereby, if node n_i belongs to a block with expected degree of k , we put the corresponding Rich Block Coefficient of this club in $NRC^{RB}(n_i)$, in the same way, if node n_i belongs to club of degree k , $NRC^{RC}(n_i)$ will be set to the respective Rich Club coefficient. Eventually, to measure similarity between the coefficients, we use coefficient of determination, (r^2).

Richness Identity

In addition to coefficients, it is also crucial to examine the similarity between significantly detected RB and RC organisations. For each simulated network, we use the RB detection method, discussed in section 4.1.3, and RC detection methods, discussed in section 3.1.9. For each network, we produce two separate Richness Identity (RI) vectors, one for the RB organisation, RI^{RB} , and another for the RC organisation, RI^{RC} . A RI vector is a binary vector of size $1 \times N$, where each element indicates whether node n is part of a significantly

detected either RB or RC organisation. Hence, if node n belongs to a significant RC organisation, its corresponding element in RI is set to one, otherwise it is set to zero.. Similarly, if node n is a member of a block which has been detected as a significant RB organisation, then $RI^{RB}(1, n_i) = 1$, otherwise $RI^{RB}(1, n_i) = 0$. To measure the similarity between RI^{RB} and RI^{RC} we use the Adjusted Rand Index (ARI). The ARI forms a 2×2 contingency table for comparing identity RI^{RC} and RI^{RB} . It can be written as:

$$ARI(RI^{RB}, RI^{RC}) = \frac{\binom{n}{2}(a+b) - [(a+b)(a+c) + (c+d)(b+d)]}{\binom{n}{2}^2 - [(a+b)(a+c) + (c+d)(b+d)]}, \quad (4.33)$$

where a is number of objects in a pair that are placed in the same group in RI^{RC} and in the same group in RI^{RB} ; b is the number of objects that are placed in the same group in RI^{RC} but in different groups in RI^{RB} , c is number of pairs that are placed in different group in RI^{RC} but the same group in RI^{RB} , d is number of pairs that are both differently grouped in RI^{RC} and RI^{RB} . Any overlap between RI^{RC} and RI^{RB} is determined by a value in between 0 and 1.

4.3 Data

Aiming to show how a Rich Block can be used to find a highly connected network core and how it can be useful in identifying the changes in topological features of the functional connectome during a disease, we use Resting-state scans of 13 healthy subjects (age= 33.3 ± 9.2 years, male=14) and 12 schizophrenic subjects (age= 32.8 ± 9.2 years, male=10). For further details about the pre-processing and network construction procedures, please see section 3.2.

4.4 Results

4.4.1 Simulations

In this section we demonstrate the results of the simulations and validations that were discussed in section 4.2.1. First, we show the validation of the Tensor null model and compare it with Maslov & Sneppen null model. Second, we validate the Rich Block by showing how RB coefficients and organisation can estimate the RC coefficients and organisations. Finally, we validate the Degree Exceptionality by showing how likely it is that the DE detects simulated block hubs.

Tensor Null Model Validation

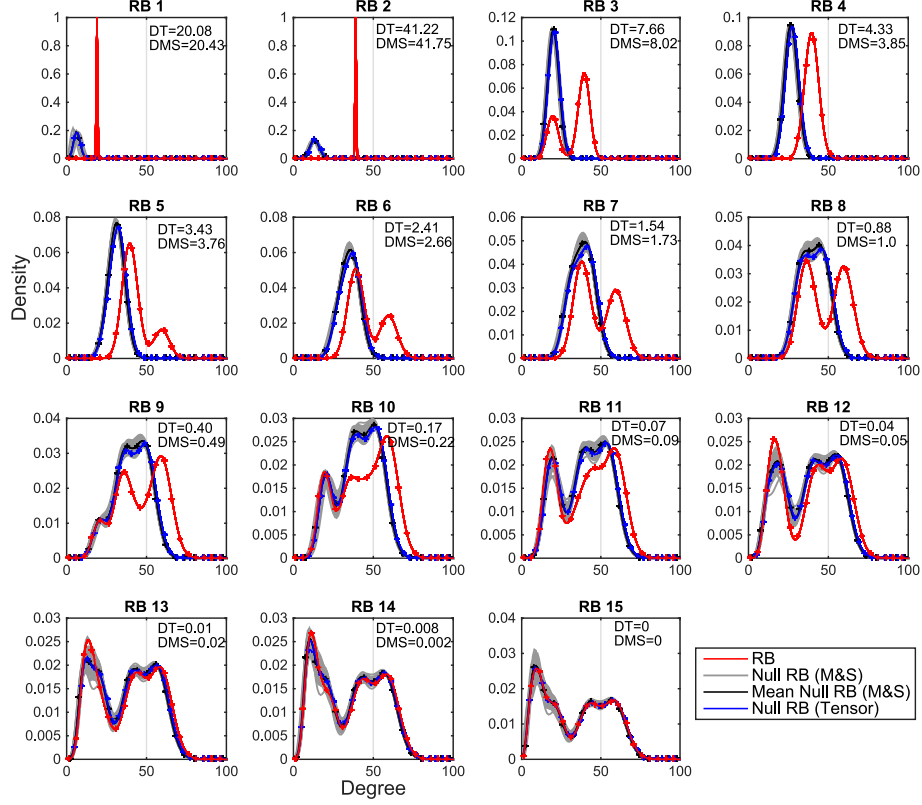


Figure 4.4: Shows probability density functions of how the proposed Tensor Null Model behaves in Rich Block organisation.

In Figure 4.4 we show the behaviour of the probability density function of the original Rich Block $f^k(\rho_i)$, Tensor Null Model $f^k_{Null-TNM}(\rho_i)$ and the Maslov & Sneppen $f^k_{Null-MS}(\rho_i)$ during formation of RB organisations of degree k . Since the network under study was constructed from 15 equal-sized blocks the number of Rich Blocks ranged from 1 to 15. It is important to note that the first sub-figure, RB1, shows the degree distributions of the diagonal mini block of the block with the highest connection rates. The last sub-figure, RB15, shows the degree distributions of the overall network. After the second RB organisation, RB2, the distance between the null models and the network decrease until RB15 where they are identical. The distances between the TNM and the estimated network are slightly lower than the corresponding distances between the MS and the estimated network, however, in RB15 the results suggest that these distances are almost zero. Also, applying the coefficient of determination between $\Pi^{Null(M\&S)}$ and $\Pi^{Null(TNM)}$ suggests 98% similarity between the two connection matrices.

Rich Block Detection and Rich Block Coefficient Validation

In Figure 4.5, each sub-figure illustrates the coefficient of determination between the Nodal Richness Coefficients and the Adjusted Random Index between Richness Identities. The results suggest the estimations, specifically in terms of Richness Identity, tend to improve upon increasing the number of nodes and blocks. For example, in a network with 300 nodes and 30 blocks is significantly better than the results suggest for a network of 60 nodes and 6 blocks. It is worth mentioning that in both cases, there are exactly 10 nodes within each block. Overall, Figure 4.5 shows that, as the number of nodes and blocks grows, the RB coefficients and organisation become more similar to the RC coefficients and organisations.

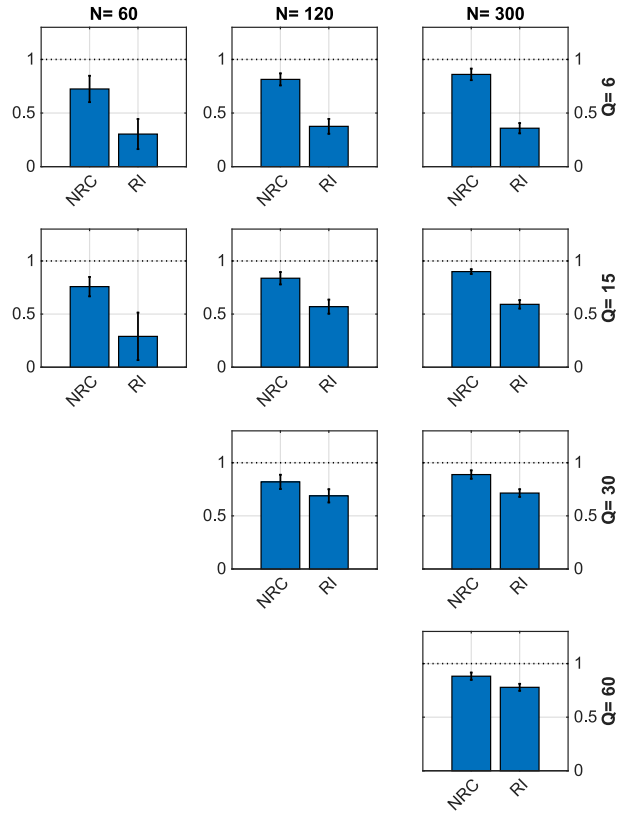


Figure 4.5: Shows results of RB estimation for both coefficient and organisation. Each column represent networks with different size and each row represent networks with different block numbers. In each subplot, the left-hand bar represent the coefficient of deterministic (r^2) between the RC and RB coefficients and right-hand bar represent the Adjusted Rand Index between the RC and RB organisations.

Degree Exceptionality validation

Figure 4.6 illustrates the behaviour of the Degree Exceptionality for different sets of networks sizes and block numbers. Each of these simulation scenarios was repeated for 1k times and the success rate was obtained as the ratio of the number of times that the DE successfully detected a hub block over total number of iterations. For networks with only one hub block, the first row of Figure 4.6, the DE could successfully detect the hub block across all iterations. However, in networks with two hub blocks, the second row of Figure 4.6, this rate reduces to $\approx 93\%$. Eventually for networks with three hub blocks, the DE algorithm could only find hub blocks across $\approx 86\%$ of the iterations.

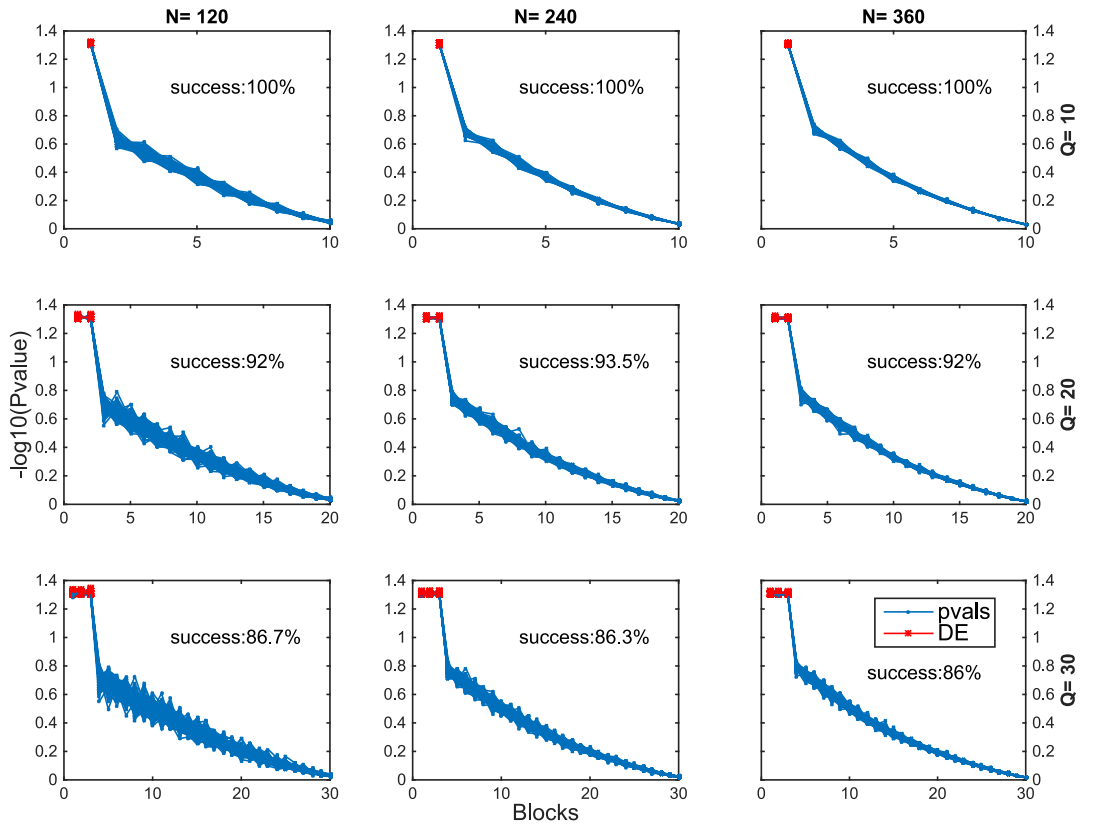


Figure 4.6: Shows results of the Degree Exceptionality on simulated networks. Each column represents networks of different sizes and each row represents networks of different block numbers. The number of hubs in each network was defined as 10% of the number of blocks.

4.4.2 Het-SBM on rs-fMRI of Healthy and Schizophrenic Subjects

Figure 4.7 shows the rearranged empirical functional connectivity of healthy and schizophrenic subjects according to the optimal block assignment obtained from Het-SBM. The results

suggest that optimal number of blocks for both groups is $Q = 31$. For a more comprehensive evaluation of results we divide the blocks into two types: coherent blocks and combinatory blocks. A block is called coherent when more than 75% of the regions inside it come from the same anatomical part of the brain. Whereas, combinatory blocks are referred to as blocks which are spatially spread across the brain and have been formed by regions of different anatomical areas. For assigning each block as either coherent or combinatory, we assume that the brain has been divided into nine separate anatomical regions: five lobes from the Cerebral cortex (Frontal lobe, Temporal lobe, Parietal lobe, Occipital lobe, Limbic lobe, Insular lobe), Thalamus, Cerebellum and Basal Ganglia.

The results suggest that 12 out of 31 blocks were identified as coherent and the 19 remaining blocks were identified as combinatory. Block 1 and Block 4 are consistently formed by regions from the middle Temporal regions. Also, block 16 was formed by regions from superior Temporal lobe. Blocks 21, 23, 26, 27 and 19 were constructed from regions mainly from the Frontal lobe and blocks 8, 17 and 20 were formed by regions from the Parietal lobe. Regions in block 6 are completely from the occipital lobe while blocks 2 and 19 were formed by regions of the cerebellum and limbic lobe respectively. Table B.1 shows the share of each anatomical region in combinatory blocks. By projecting the block assignments on a brain surface (see Figure B.1) we also realise that although the functional connectome lacks any form of anatomical information, the Het-SBM could capture the spatial information reasonably well, which suggests that the pattern of the forming connections in anatomical regions are consistent with their function.

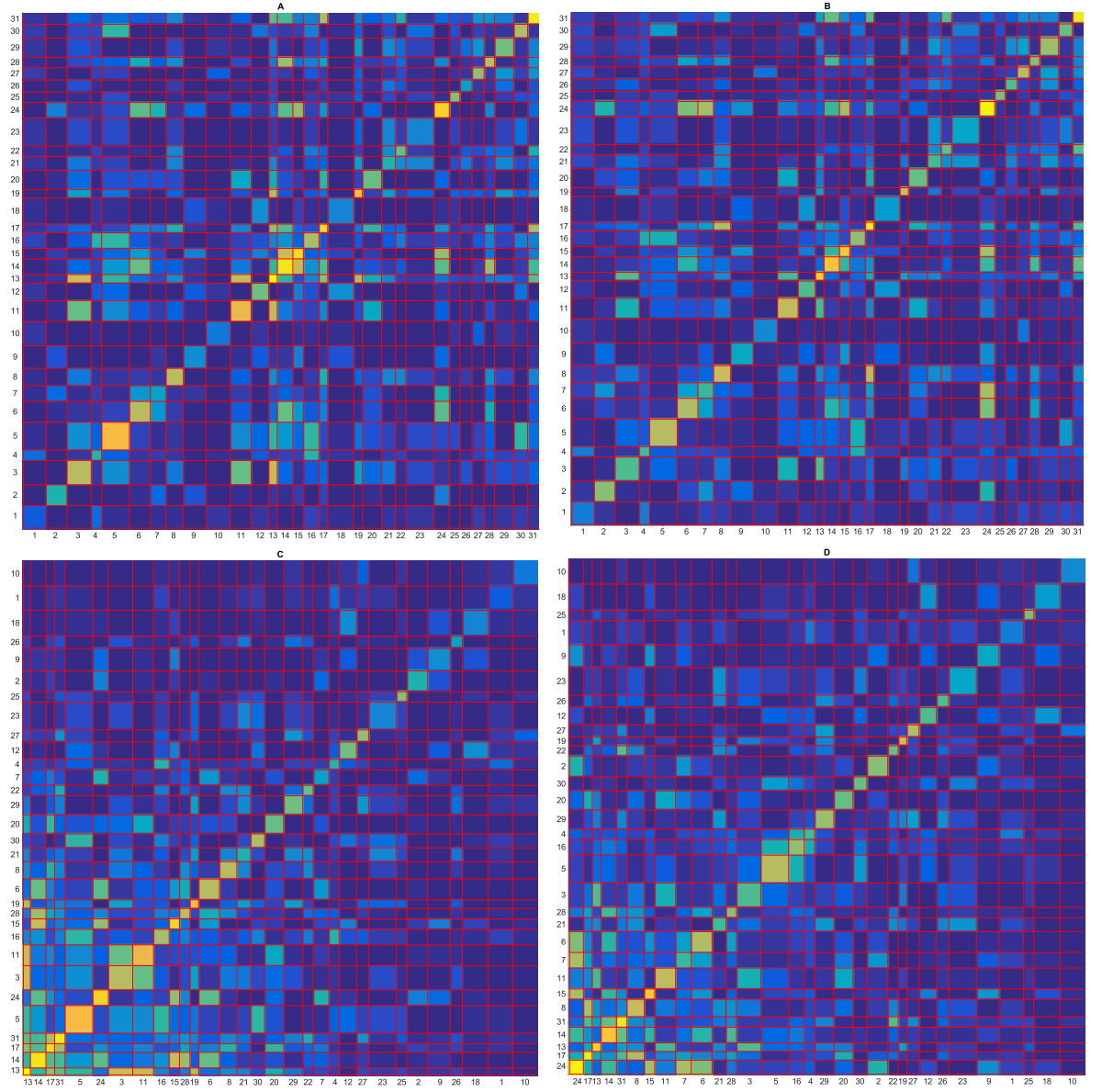


Figure 4.7: Connection probability matrix of control **A** and Schizophrenic patients **B**. **C** and **D** are sorted connection probability matrices based on the decreasing expected degree of each block.

4.4.3 A Rich Block for rs-fMRI Schizophrenia

Rich Block Organisations

Figure 4.9 shows the RBI of each group of healthy (top panel) and schizophrenia (bottom panel) subjects. Each RBI shows blocks that were identified as a significant Rich Block organisation. The results of within group analysis suggest that blocks 13, 14, 17, 31 were

identified as significant RB organisations, while, in schizophrenia only blocks 17 and 24 were identified as RB organisations. Hypothesis testing between $RBI_{Control}$ and $RBI_{Patient}$ reveals that Blocks 13, 14, 31 are no longer part of the detected Rich Block organisations in schizophrenic functional connectivity, however, a new block, 24, joins the RB organisation during schizophrenia. Therefore, the only block that remains unchanged during the disease is block 17. By referring to Table B.1, we identify blocks 13, 14, 24 and 31 as combinatory blocks. These blocks were spread across the Parietal, Limbic and Frontal lobe. Block 24 which is also combinatory is formed by regions from the Occipital Lobe and the Cerebellum. Finally, block 17 is the only block which has appeared to be among coherent blocks and it only has regions from the Parietal Lobe.

For further visual examinations, if a block was identified as a significant RB organisation in more than 25% of the individuals within a group then we consider that block as part of the overall RB organisation of a group. Figure 4.8 shows the configuration between blocks of the RB organisation and the rest of the network in both healthy and schizophrenia subjects.

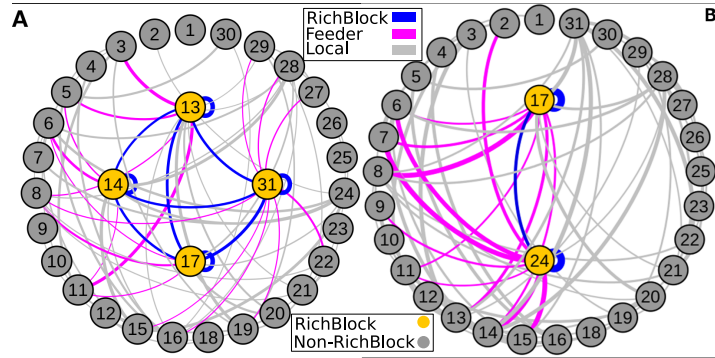


Figure 4.8: illustrates a circular layout of the group-averaged probability connection matrix, Π , of the controls **A** and patients **B**. Detected RBs are located in the middle of the layout. For the sake of visualisation, we only drew the top 30% of the strongest connections.

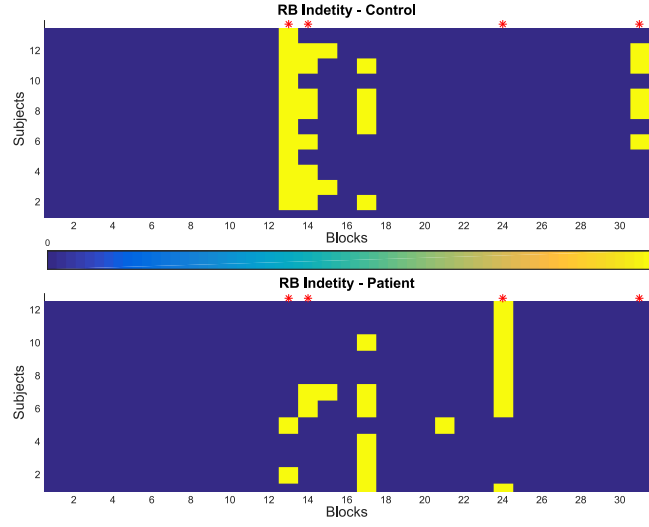


Figure 4.9: Shows the Rich Block Identity (RBI) of controls (top) and patients (bottom). The yellow cells indicate that block q for subject s is a significant Rich Block. Red asterisk indicates statistical differences between two RBIs.

Rich Block Coefficient

Figure 4.10 A and C shows BRBC of the controls and patients regardless of whether they were detected as significant RBs. In the same figure, sub-figures B and D show results of statistical inference between the mean of each block (column) of $BRBC^{control}$ and $BRBC^{patient}$: Blocks 1, 2, 4, 6, 7, 8, 9, 17, 21, 23, 24, 26 and 29 have relatively larger normalised RB coefficients in schizophrenic subjects, whereas, blocks 5, 18, 19 in the control have larger normalised RB coefficients in healthy subjects. 69% of the blocks which were identified as larger in schizophrenia are among the coherent blocks, while, 66% of the blocks with higher normalised RBs in healthy subjects are among the combinatory blocks.

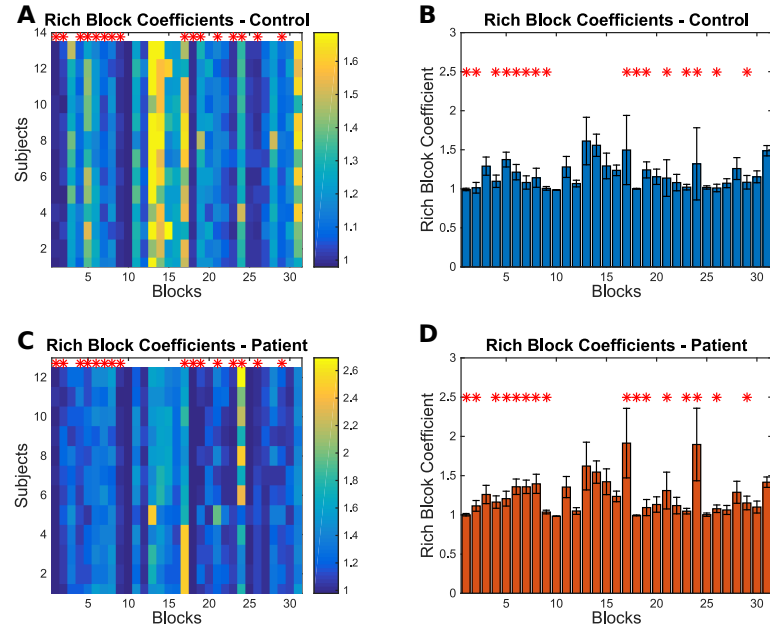


Figure 4.10: A and C shows Block-wise Rich Block Coefficient (BRBC) in control and patient respectively. B and D shows mean of each column in A and C. Red asterisk shows statistical significance between mean of RB values of two groups.

In terms of anatomical characteristics of the blocks which are significantly higher in schizophrenia, blocks 2, 7, 9, 24 which were found to have a larger RB coefficients, are among the combinatory blocks which highly overlap with the Cerebellum. Blocks 21, 23, 26 and 29 are among the blocks which highly overlap with the Frontal Lobe and, finally, block 8 and 7 are among the coherent blocks of the Parietal Lobe. On the other hand, among the blocks which experience an increase in their RB coefficients, blocks 5 and 18 have combinatory blocks which are larger overlaps with Frontal, Parietal and Limbic lobes, whereas, block 19 is a coherent block of the Limbic system.

Local and Feeder Blocks

Figure 4.11 shows the density of each class of connection (i.e. RB, Feeder and Local). Figure 4.11.B suggests that the connection rate among blocks of the RB organisation of the functional connectome reduces during schizophrenia. Figure 4.11.C suggests that the densities of feeder connections are also reduced during the disease. However, Figure 4.11.D suggests that the schizophrenic functional connectome enjoys a larger density of local connections in comparison to the controls, while, the density of the local connections in healthy individuals as a large variance. Considering that we used the same overall threshold across the subjects, the results suggest that during schizophrenia the rate of connection is shifted

from the RB organisation to local connections.

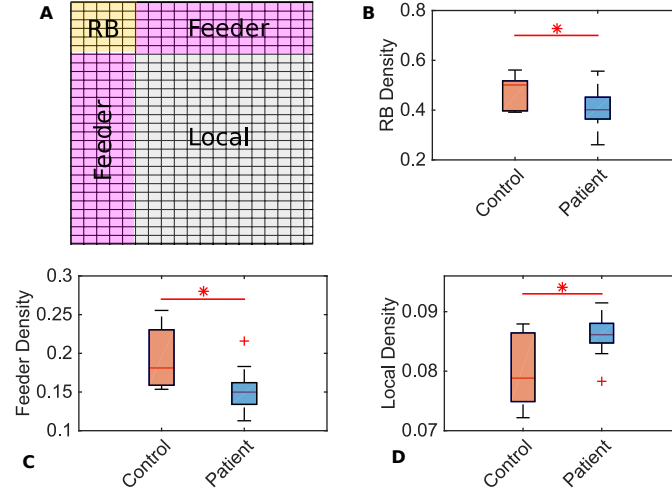


Figure 4.11: **A.** illustrates the organisation of the local and feeder connections along with RB in a synthetic adjacency matrix. **B.** Illustrates the difference in density of the RB connection individuals in healthy and schizophrenia. **C.** Shows the feeder density of individuals in healthy and schizophrenia subject groups. **D.** Shows the local density between individuals of healthy and schizophrenia. The red asterisk indicates the statistical difference.

4.4.4 Degree Analysis on rsfMRI Schizophrenia

Figure 4.12 A and C show binary matrices representing the DE blocks in individuals across the healthy and schizophrenic groups, respectively. In the healthy group, block 13 ($\lambda_{13} = 74.2$) was detected as the DE unanimously over all individuals, whereas, in the schizophrenia blocks 17 and 24 was identified as a hub block. It is notable that block 17 ($\lambda_{17} = 55.43$) and 24 ($\lambda_{24} = 59.80$) are not consistent across individuals as in only one subject both of these blocks were identified as hub blocks. Figure 4.12 B and D show the expected degree, λ , of each block across subjects of the healthy and schizophrenic groups.

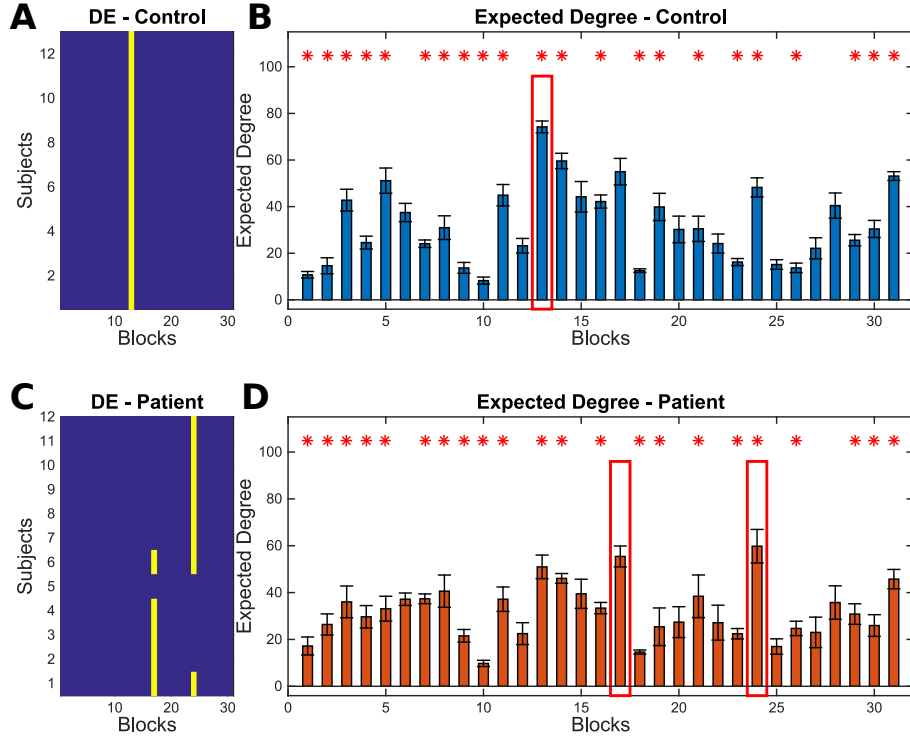


Figure 4.12: **A** and **B** are binary matrices of the DE effect in healthy and schizophrenia groups. **B** and **D** are the expected degree of each block for the individuals in the control and patient group. Red asterisk indicates the statistical difference and red rectangles indicate the degree exceptional blocks.

Figures 4.12.B and 4.12.D show the expected degree of blocks in healthy and schizophrenia subjects, respectively. Statistical testing between the two groups suggests that only block 24 is among the hub blocks that experience a decrease during the disease. Further investigation of the results suggests that blocks 3, 5, 11, 13, 14, 16, 19, 30, 31 are significantly larger for the healthy subjects. It is notable that blocks 5 and 19 were also identified to have significantly larger RB coefficients. On the other hand, blocks 1, 2, 4, 7, 8, 9, 10, 18, 21, 23, 24, 26, 29 enjoy higher expected degree in schizophrenia. Apart from blocks 10 and 18, the remaining blocks were also identified to have higher RB coefficients. In terms of the anatomical characteristics of the blocks, 77% of those blocks that were identified to be larger in schizophrenia are among the coherent blocks, whereas, 78% of the blocks that were shown to be larger in the healthy group are among the combinatory blocks.

We also examined the degree exceptionality on group-level connection rate matrices directly obtained from Het-SBM. The results suggest that block 13 ($\lambda_{13} = 72.44$) in healthy subjects and blocks 17 ($\lambda_{17} = 54.38$) and 24 ($\lambda_{24} = 58.30$) in schizophrenia are hub blocks.

For full analysis of the DE on the group-level connection rate matrices see Appendix B.6.

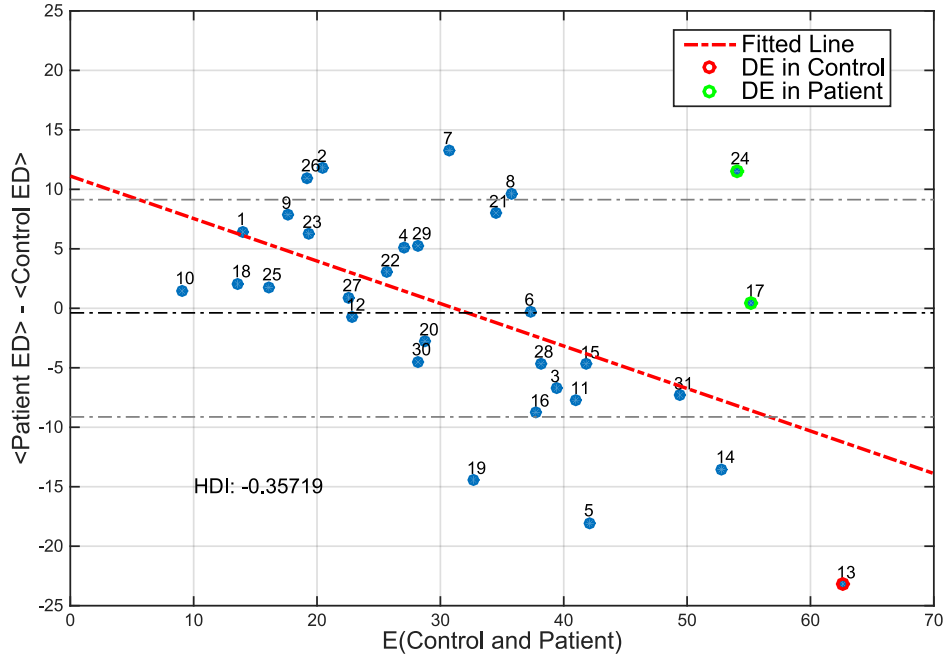


Figure 4.13: Illustrates a Bland-Altman plot to investigate hub distribution between healthy and schizophrenic patients. The black dotted line shows the mean of the differences. The grey dotted lines indicate the 1 SD distance from the mean.

Figure 4.13 illustrates the Bland-Altman (Hub Disruption) between the expected degree of blocks in the healthy and schizophrenia groups. The slope of the fitted line to changes is -0.35 which suggests that the blocks with a higher rate of connection have a less prominent role during Schizophrenia.⁷ The block that has mostly experienced this disruption is block 13 which is a combinatory block of the Frontal lobe and the Parietal lobe. This block was identified as the degree exceptional and the member of the RB organisation in the healthy group. Block 14 which is combinatory block of the Parietal lobe, Occipital lobe and Limbic lobe is also among the healthy RB organisation which experiences a sharp decrease during Schizophrenia. Block 19 which is a coherent block of the limbic lobe and block 5 which is combinatory block of the Frontal lobe, Temporal lobe and Parietal lobe are also among the highly connected blocks which experienced a disruption during the disease. On the other hand, blocks that gained connectivity during schizophrenia are blocks 7 and 24 which are combinatory blocks of the Occipital lobe and the Cerebellum. Both of these blocks were identified as the DE and member of RB organisation in the schizophrenia

⁷In this section we only report changes in degree of blocks whom locate above or below 1σ of the all changes.

group. Block 2 which is a coherent block of the Cerebellum, blocks 26 and 21 which are coherent blocks of Frontal lobe and block 8 which is a coherent block of the Parietal lobe are also among the less connected blocks which experienced an increase during schizophrenia. It is notable that block 17 is a member of RB organisation in both healthy and schizophrenia groups, suggesting almost zero changes in its expected degree during schizophrenia, despite the fact, it was identified as a hub block by the DE.

4.5 Discussion

In this chapter, we proposed a novel graph theoretical measure called the Rich Block. the Rich Block is an improvement over the conventional Rich Club in the sense that it facilitates the group inference between subjects, in contrast to the Rich Club which was shown to be limited in group analysis of neuroscientific datasets. Further to the group inference, the Rich Club forms the so called 'clubs' that are merely based on the degree of nodes, whereas, the Rich Block forms the Rich Block organisations (former clubs) based on block assignments estimated by a Stochastic Block Model. This results in having nodes grouped not only based on their degree, but also based on how the nodes interacting with rest of the network. For example, if a stack of nodes has exactly the same degree, but different inter- and intra- connectivity pattern, they will be grouped as two separated blocks. It is also worth noting that with the proposed Tensor Null Model, the process of detecting the Rich Block phenomenon becomes remarkably fast.

Our findings suggest that the Rich Block organisation in functional connectivity of schizophrenia is impaired especially in the Frontal and Parietal lobes. In healthy subjects, the Rich Block organisation is formed by four blocks which all are involved in the DMN of the human brain. Three of these blocks are combinatory blocks which are mainly formed by a collection of nodes from the Parietal, Frontal and Limbic systems and one remaining block is a coherent block of the Parietal Lobe. These blocks include regions from the Posterior Cingulate, Precuneus and Paracentral Lobule. However, our findings suggest that in the schizophrenic functional connectome, two combinatory blocks of the DMN no longer become members of the RB organisation, Instead, a combinatory block of the Occipital and Cerebellum joins the RB organisation. It is worth stressing that RB coefficients of the RB organisations are not different apart from block 17 that suggests an increase during the disease. Interestingly, the expected degree of this block is quite similar across groups. This result suggests that although block 17 retained its RB organisation membership, as well as its expected degree during schizophrenia, it was affected by changes in the topological features of the other members of the RB organisation.

The majority of blocks (70%) that were found to have higher RB coefficients in

healthy subjects are among the combinatory blocks. These combinatory blocks are formed by spatially distinct brain regions, including the Precentral Gyrus of the Frontal Lobe, the Postcentral Gyrus of the Parietal Lobe, the Cingulate Gyrus of the limbic system and Insular Lobe. On the other hand, the same proportion of blocks which were found to have higher RB coefficients during schizophrenia are among coherent blocks including the Cerebellum, the Occipital lobe, medial and superior gyrus of the Frontal lobe and middle gyrus of the Temporal lobe. The realisation that a large proportion of the RB organisation in healthy subjects is formed by combinatory blocks may emphasise the role of the RB organisation as a backbone to the functional connectome as it was previously shown that the RC organisations play a similar role within the human connectome (de Reus and van den Heuvel, 2014). However, this role is later diluted during the disease as only one combinatory block, 24, remains as part of the schizophrenic RB organisation.

We also examined three higher level connections: connections within the RB organisation (RB connections), connections between the RB organisation and the rest of the network (feeder connections) and connections between the non-RB parts of the network (local connection) across the healthy and schizophrenic subjects. The results indicate that the density of connections within the RB organisation reduced during schizophrenia. The reduction in the RB connection of the schizophrenic connectome had been previously reported in the functional (Yu *et al.*, 2013b) and the structural (van den Heuvel *et al.*, 2013b) connectome. Additionally, the results suggest a reduction in feeder connections of schizophrenic patients, whereas, density of local connections increase during the disease. Bearing in mind that density of all subjects was fixed via defining a global density, these results suggest that the density of connections demonstrates a shift from the RB organisation to a local level. The shift in density of the connections from the RC organisation to the local connection level was previously reported in the structural connectivity of schizophrenic patients (van den Heuvel *et al.*, 2013b; Collin *et al.*, 2014b).

We also proposed a method that aims to identify the hub blocks, called the Degree Exceptionality. Although hubs have a fairly straightforward definition, identifying them is rather an arbitrary decision since it is unknown that above what threshold a node (or block) can be considered as a ‘hub’. The Degree Exceptionality simplifies the process of finding hubs by detecting the block that has a remarkably high inter- and intra- connection rate. The likelihood of a block to be identified as a hub block is heavily associated with the size of the block and its connection rate. Our findings in the healthy human brain suggest that a combinatory block of the DMN which includes the Cingulate Gyrus, Paracentral Lobule is a hub block of the healthy human brain. The same approach identifies two blocks as a hub in the schizophrenic brain network; one is a coherent block of Parietal lobe which includes the Precuneus of the DMN and the other is a combinatory block of the Cerebellum and

the Occipital lobes. Although it is notable that the hub block found in the healthy subjects suggests a remarkably higher rate of connection compared to two hub blocks found in the schizophrenia. This findings suggest that the schizophrenic functional connectivity not only demonstrates a reduction in the strength of hubs, but also suggest a shift from DMN-centred hubs to not-Parietal, not-Frontal hubs such as the Occipital and the Cerebellum. It is also worth noting from DE results in healthy subjects, that the hub block is a combinatory block of the DMN consistently over all subjects, whereas, in the schizophrenia the hub blocks include a combinatory blocks of the Occipital and the Cerebellum and coherent block of the Parietal lobe, however, only one schizophrenic patient (subject 6) demonstrate the both hub blocks, simultaneously. It is worth mentioning that less prominent role of hubs (Lynall *et al.*, 2010; Drakesmith *et al.*, 2015) as well as emergence of the non-Parietal hubs (Bassett *et al.*, 2008; van den Heuvel *et al.*, 2013b; Collin *et al.*, 2014b) has already been reported.

The changes in the expected degree of blocks was also summarised using a Bland-Altman plot as an adopted version of the Hub Disruption notion which was initially proposed by Achard *et al.* (2012). This plot suggests that blocks which represent the DMN (such as Blocks 13 and 14) which are combinartory blocks of the Frontal (Paracentral Lobule), the Parietal (Precuneus) and the Limbic (Cingulate Gyrus) lobe along with other combinatory of block of insular and coherent block of limbic lobe demonstrated the largest disruption during schizophrenia, whereas, other parts of the Parietal lobe along with the Cerebellum and the Occipital lobes gained connection during schizophrenia.

Finally, it is important to note that the findings presented in this chapter regarding the changes in functional connectivity in schizophrenia using a Rich Block have a strong correlation with the findings from the same data-set using a Rich Club which were reported in section 3. First, in chapter 3, the results suggest that during schizophrenia, the hub nodes lose their prominent role by experiencing a reduction in their connection rate. The hub nodes which have drastically experienced this reduction are among the DMN members of the Parietal Lobe such as the Precuneus, the Cingulate cortex and the Paracentral regions. On the other hand, the non-Parietal/non-Frontal regions such as the Cerebellum and the Occipital regions demonstrate an increase in their connection rate. Similar results have been found by using the DE algorithm and a Disruption Plot as in this chapter. Second, investigating the changes in the RC organisations in schizophrenia, discussed in section 3.3.7, suggests that several nodes of the the Parietal and the Frontal lobes no longer contribute to the RC organisation. Similarly, in this chapter, the results suggest that the RB organisation in healthy individuals is formed by blocks 13, 14, 17 and 31 which are mainly parts of the Frontal and Parietal lobes however, in schizophrenic patients the RB organisation was merely detected in blocks 17 and 24. Block 24 is formed by parts of the Cerebellum and the Occipital lobes.

4.6 Summary

In this chapter we presented a series of stochastic measures which help the network neuroscientists to gain a better understanding regarding the cores of the human functional connectome.

- Previously proposed Het-SBM method was applied to the same data-set as Chapter 3, schizophrenic data-sets, and showed how the common block assignments can facilitates the group inference.
- A novel stochastic method was proposed, called the Rich Block, which outperforms the Rich Club in terms of three main factors: Firstly, the RB facilitates the group inference as it employs Het-SBM in estimation of the block assignments. Secondly, the Rich Block takes into account the effect of covariates, and therefore, can deliver a much robust statistical inference. Thirdly, the Rich Block is a complimentary method to the stochastic block modelling, which considers the pattern of connectivity as well as the degree of nodes in identifying the connectome cores.
- Similar to the RC, the RB also requires a randomisation method (or a null model) aiming to ensure that the effect identified as core is a genuine effect. A randomisation method, called the Tensor Null Model, was proposed which is novel in stochastic domain and speeds up the the conventional, time-consuming, randomisation process by 63k times.
- A novel stochastic method was proposed, called the Degree Exceptionality. The DE facilitates the process of finding hub blocks. In previous literature, hubs are conventionally selected by arbitrary measure, however, the DE can detect the nodes which enjoy high number of connection and the same pattern of connectivity.
- An improved version of the Bland-Altman plot, or conventionally called the disruption plot, also presented in this chapter which can be used in summarising the changes during a disorder. We used the disruption plots in Chapter 3 to summarise the changes in degree, whereas, in this chapter they are used to summarise the change in expected degree of blocks.
- Finally, we use simulations as well as real data to validate the proposed methods. We also demonstrated how they can help the researchers to detect the anomalies in functional connectome due to psychiatric disorders.

Chapter 5

Impact of Autocorrelation on Topological Features of Functional Connectivity of the Human Brain

Functional brain imaging methods allow the examination of interactions between the brain regions through temporal similarity measurement. Pearson's correlation, as a well-known temporal similarity measure, determines the interaction between two brain regions as a degree of the linear relationship between two BOLD signals (Smith *et al.*, 2011b). However, the temporal similarity between two brain regions can be produced by nuisance effects rather than physiologically meaningful dependencies. Correlated (non-white) noise, can result in a single region's BOLD signal being correlated with lagged versions of itself (Weisskoff *et al.*, 1993). Correlated noises in BOLD signals have several sources including cardiac and respiration processes, instrument instability and even fMRI pre-processing steps such as registration (Lund *et al.*, 2006). Signals that exhibit this property are conventionally called *autocorrelated signals*. The presence of temporal autocorrelation may bias the estimated correlation coefficient between a pair of regions. Further, the estimated uncertainty in the correlation estimate can be biased, as reflected in the actual degrees of freedom (DF) being lower if the noise was white. Furthermore, there is also evidence that the sampling variability in Pearson's correlation can, by itself, induce non-trivial topological effects such as non-biological modularity and small-worldness (Zalesky *et al.*, 2012a). Whether biasing the correlation estimate itself or disturbing the DF, temporal autocorrelation can result in inflated false positive rates.

While the various network statistics reviewed in Chapter 3 can be defined for weighted graphs, they are most commonly applied to binary graphs. There are two general approaches to thresholding continuous correlations into binary edges: [1] Statistical Thresh-

olding (ST), which keeps any edge that rejects the null hypothesis of zero correlation at a given critical threshold α (Bassett *et al.*, 2011; Fornito *et al.*, 2016); [2] Density Thresholding (DT), which uses the threshold necessary to obtain a desired edge density. However, as the exact choice of density is arbitrary, the graph metrics are normally examined over a range of different density thresholds. Crucially, in multi-subject studies of functional connectivity, all subjects have the same density (Bullmore and Bassett, 2011; Fornito *et al.*, 2016).

These two methods of thresholding, ST and DT, can be largely affected by the presence of the aforementioned autocorrelation in each time series. As in ST, inflated degree-of-freedom due to the presence of autocorrelated time series can lead the statistical inferences being liberal and produce excess false-positives. False positives can have a particular impact on descriptors of the topology of the human brain network (Zalesky *et al.*, 2016), thus motivating ST with stringent control of Type I error risk. While, in DT, there can still be a problem because the variation in temporal autocorrelation over nodes will result in the correlation coefficient having different sampling distributions for different pairs of nodes. The end result is that, while a fixed threshold is used to obtain a given degree, the actual meaning of that threshold varies over edges. There are also two recent studies which examined the effect of inflated degree of freedom in the thresholding of functional connectivity. One study proposes an unbiased estimator of sample correlation variance for an excessively oversimplified first-order autoregressive model of BOLD signals (Arbabshirani *et al.*, 2014), while the others propose a wavelet-based method using a whitening property of wavelets to estimate the effective degree of freedom in rarely-popular wavelet transformed correlation coefficients (Patel and Bullmore, 2015).

The term degrees of freedom can have different meanings depending on context, and so we pause for some definitions. In linear models, "degrees of freedom" typically refers to the error DF, and is the number of observations (N) minus the number of predictors. It is perhaps best known for its role in estimation of the error variance. The naive estimate of the error variance, the residual mean squared¹, is a biased estimate for the true error variance, being off by a factor of DF/N . Thus the RMS needs to be scaled by a correction factor of N/DF or, equivalently, an unbiased estimator is found as the residual sum of squares divided by DF. This unbiased variance estimate is crucial for all linear model inference, in particular assuring that T-tests and F-tests will follow their nominal distributions.

In functional connectivity, Fisher's transformation is applied to the correlation coefficient, giving a measure with approximate normality and constant variance (details in Section 5.1.1 below). Under the usual assumptions of independence, this variance is $1/(N - 3)$, but autocorrelation can invalidate this result. Typically, temporal autocorrelation reduces

¹Residual mean squared (RMS) is the sum of squared residuals divided by N

the information content in the time series, and leads to *more* variable estimates of correlation between two variables. That is, if a Fisher-transformed correlation was naively scaled by $\sqrt{N-3}$, the resulting Z statistic would have variance greater than unity and lead to an inflated false positive risk.

The focus of this chapter is an adjustments to Pearson’s correlation such that the transformed statistic has nominal variance. As detailed below (Section 5.1.2), we seek a correction factor (CF) such that scaling an estimate of correlation (before or after Fisher’s transformation) by a value of $\sqrt{N/CF}$ produces a measure with unit variance. CF will be approximately 1 for independent data, but will increase with the severity of the temporal autocorrelation. The effective degrees of freedom can be regarded as N/CF .

While a relatively recent concern in the analysis of the functional connectivity, the problem of inflated error degrees of freedom in the analysis of first-level GLM of fMRI data has a long history. In the late 1990’s the fMRI analysis strategy advocated by Karl Friston and colleagues was to smooth BOLD time series data using a high-pass filter. While this is not statistically optimal (it is ”colouring” instead of the optimal ”whitening”), it precluded the need to model the intrinsic temporal autocorrelation, and allowed for precise estimates of the degrees of freedom (Friston *et al.*, 1995). Subsequently Friston’s group dropped temporal smoothing, and adopted a whitening approach where the autocorrelation is estimated globally over the brain (Friston *et al.*, 2000); others proposed estimation of autocorrelation individually at each voxel (Bullmore *et al.*, 2001) or voxel-wise with spatial regularisation (Woolrich *et al.*, 2001).

Long before functional neuroimaging existed this issue was taken up in economics (Hurwicz, 1944; Bartlett, 1946, 1935). In his seminal 1935 paper, Bartlett introduced a formulation for estimating the sample variance of the correlation coefficient using a first-order autoregressive (AR(1)) time series model. A decade later, he produced a more general result for AR models of any order. Several neuroimaging studies have used both AR(1) and more general versions of the Bartlett’s DF correction however, extensive reading of the neuroimaging literature finds that applications of Bartlett’s corrections are inconsistent, and at best are poorly described. For example, for a DF correction, many papers reference Fox *et al.* (2005) that opaquely cites a Bartlett correction in a lengthy textbook. The exact form of this is described only in prose in Van Dijk *et al.* (2010) which ignores the impracticality of estimating the true (unobserved) autocorrelations. In addition, the method for correcting DF used throughout the literature is mainly global and consequently fails to address the heterogeneity of the effect of autocorrelation (Golestani and Goodyear, 2011; Johnston *et al.*, 2008; Goparaju *et al.*, 2014; Hale *et al.*, 2016).

In this study we introduce a bivariate form of Bartlett’s correction, accounting for the possibly different temporal autocorrelation in each of the two time series considered.

Re-formulating Bartlett’s correction factor for correlation coefficients of bivariate time series allows us to calculate the autocorrelation effect heterogeneously. In other words, instead of considering the Bartlett Correction Factor (BCF) as a global adjustment for each subject or each study, we rather calculate this correction factor for every edge without assuming identical autocorrelation at each node. Such a bivariate, heterogeneous correction factor should make both traditional statistical and density thresholding methods more robust against false-positives, and in particular ensure that the false positive risk is the same for all edges. We validate the proposed method with real-data and compare it with existing methods. We also examine how the choice of pre-processing steps can affect this estimation and, essentially, how these corrections can affect the quantification of topological properties of functional connectivity. To investigate the effect of this correction in graph theoretical methods, we examine the statistics on three fundamental features of the functional connectome: Integration, segregation and complexity. In addition, we validate the proposed method by four separate cohorts from different acquisition sites to ensure that the results are reproducible.

5.1 Methods

5.1.1 Functional Connectivity as a Correlation Matrix

To form a functional connectivity of each subject, we use Pearson’s correlation to measure the temporal similarity between each pair of the time series, x and y , representing a brain region. Therefore, each ROI is considered as a node and the correlation coefficient between x and y is considered as the weight to the edge formed between the two nodes. The population (‘true’) correlation coefficient, ρ_{xy} , between two random variables X and Y is

$$\rho_{xy} = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)\text{var}(Y)}}. \quad (5.1)$$

For samples of X , x_1, x_2, \dots, x_N , and of Y , y_1, y_2, \dots, y_N , the Pearson’s correlation coefficient is an estimate of ρ_{xy} ,

$$r_{xy} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}}, \quad (5.2)$$

where N , exceptionally to this chapter, is the number of observations and \bar{x} and \bar{y} are the sample means of X and Y , respectively. This estimate is only approximately unbiased

$$E(r_{xy}) \approx \rho_{xy} - \frac{\rho_{xy}(1 - \rho_{xy}^2)}{2N}, \quad (5.3)$$

and has variance that varies with the true correlation

$$\text{var}(r_{xy}) \approx \frac{(1 - \rho_{xy}^2)^2}{N} \left(1 + \frac{11\rho_{xy}^2}{2N} \right). \quad (5.4)$$

In addition to the variance being dependent on the true ρ_{xy} , the sampling distribution can be highly non-normal. To address this, Fisher's transformation is typically applied as following

$$\hat{z}_{xy} = \text{arctanh}(r_{xy}) = \frac{1}{2} \ln \left[\frac{1 + r_{xy}}{1 - r_{xy}} \right]. \quad (5.5)$$

When the original data comprise independent, identical draws from a bivariate normal distribution, Fisher's \hat{z}_{xy} has an approximate normal distribution with mean $\text{arctanh}(\rho_{xy})$ and variance $1/(N - 3)$. Thus a standardised statistic, a Z-score, can be computed as $z = \hat{z} \sqrt{N - 3}$.

5.1.2 Correction Factor

When the independence assumption (over the N observations) does not hold, the variance of \hat{z}_{xy} will typically be inflated above $1/(N - 3)$. The usual assumption of the independence, as described, refers to a realisation that a n-lagged version of a time series has a zero correlation with the 0-lag version of the time series. We define the *correction factor* β that defines the effective degrees of freedom (EDF) $\hat{N} = \beta N$, such that

$$\text{var}(\hat{z}_{xy}) = \frac{1}{\hat{N}}. \quad (5.6)$$

The correction factor gives the EDF such that $z = \hat{z} \sqrt{\hat{N}}$ has unit variance, and thus can be interpreted as a z-score.

The correction factors that we investigate in this section can be grouped into two major categories. *Global* correction factors, that estimate a common EDF for the entire brain. *Local* corrections factors, in contrast, estimate an EDF for each pair of ROIs (or voxels) and thus account for the spatially heterogeneous effect of autocorrelation. Further, we will consider different types of local corrections, one that is *bivariate*, and allows autocorrelation to vary between the two nodes considered, or non-bivariate, which assumes the autocorrelation is the same at both nodes.

5.1.3 Naive Correction

As noted above, when the standard assumptions hold, the $\text{EDF} = N - 3$ (Fisher, 1915). Throughout this chapter, we call this method the Naive Correction as this work was moti-

vated by inflated false-positive rates found when using this approach in typical connectome analyses.

5.1.4 Autoregression and Autocorrelation

To account for the dependence seen in the time series measurements used in fMRI, we consider Autoregressive (AR) Models (Jenkins and Reinsel, 1976). A first-order Autoregressive model (AR(1)) of the time series x can be defined as

$$X_{i+1} = \Phi_1 X_i + \epsilon_{i+1} \quad (5.7)$$

for $i = 1, \dots, N - 1$, where Φ_1 is the AR(1) coefficient, and the ϵ_i are independent and identically distributed normal random variables. Using least squares, Φ_1 can be estimated as

$$\hat{\Phi}_1^x = \frac{\sum_{i=1}^{N-1} x_i x_{i+1}}{\sum_{i=1}^{N-1} x_i^2}. \quad (5.8)$$

The Autocorrelation Function (ACF), λ_s , of a stationary timeseries x is the correlation between measurements separated by s observations (Jenkins and Reinsel, 1976). The autocorrelation coefficient at lag s is

$$\lambda_s^X = \frac{\text{cov}(X_i, X_{i+s})}{\text{var}(X_i, X_{i+s})} \quad (5.9)$$

which is estimated as

$$\hat{\lambda}_s^x = \frac{\sum_{i=1}^{N-s} (x_i - \bar{x})(x_{i+s} - \bar{x})}{\sum_{i=1}^N (x_i - \bar{x})^2}. \quad (5.10)$$

It is worth mentioning that by following Eq. 5.8 and Eq. 5.10 we conclude that the first-order autoregressive coefficients are equal to the first-lag autocorrelation coefficients.

5.1.5 Bartlett's Correction - Markov Process

The earliest work on this problem comes from 1935, where Bartlett considered a number of issues related to inference on individual or pairs of AR(1) models (Bartlett, 1935). He considered the pairwise correlation, r_{xy} , between AR(1)-correlated time series models X and Y , with AR(1) coefficients Φ_1^X and Φ_1^Y , respectively. For the case when the two series are not correlated, $\rho_{xy} = 0$, Bartlett derived the variance of r_{xy} as

$$\text{var}(r_{xy}) \approx \frac{1}{N} \frac{1 + \hat{\Phi}_1^x \hat{\Phi}_1^y}{1 - \hat{\Phi}_1^x \hat{\Phi}_1^y}. \quad (5.11)$$

As per Eq. 5.6 the EDF, \hat{N}^{xy} , is the reciprocal of the variance,

$$\hat{N}_{HetBiv-AR(1)}^{xy} \approx N \times \frac{1 - \hat{\Phi}_1^x \hat{\Phi}_1^y}{1 + \hat{\Phi}_1^x \hat{\Phi}_1^y}, \quad (5.12)$$

and the correction factor is comprised of the second term in this expression.

This result is necessarily an approximation and may have degraded accuracy if ρ_{xy} is large in absolute value. Hence, [Arbabshirani et al. \(2014\)](#) suggested modifying the sampling variance, replacing the $1/N$ term with the analogous leading term from Eq. 5.4:

$$\text{var}(r_{xy}) \approx \frac{(1 - \rho_{xy}^2)^2 (1 + \hat{\Phi}_1^x \hat{\Phi}_1^y)}{N (1 - \hat{\Phi}_1^x \hat{\Phi}_1^y)}. \quad (5.13)$$

Again, the EDF for this pair of nodes is then the reciprocal of the variance:

$$\hat{N}_{HetBiv-AR(1)^{BC}}^{xy} \approx \frac{N(1 - \hat{\Phi}_1^x \hat{\Phi}_1^y)}{(1 - \rho_{xy}^2)(1 + \hat{\Phi}_1^x \hat{\Phi}_1^y)}. \quad (5.14)$$

5.1.6 Bartlett's Correction - General Process

All of the results in the preceding section are based on the assumption that each of the two time series follows an AR(1) process. However, BOLD time series, are of a far more complex temporal structure. A decade after his first work on this topic, [Bartlett \(1946\)](#) introduced results for a General Process, specified by its ACF λ_k . Unfortunately, in that work he does not consider the sampling distribution of a correlation between two time series. He does, however, introduce the notion of effective degrees of freedom. The variance of the sample ACF at lag s , $\hat{\lambda}_s$, is given as

$$\text{var}(\hat{\lambda}_s) = \frac{1}{N} \sum_{s'=-\infty}^{\infty} \lambda_{s'}^2, \quad (5.15)$$

under the assumption that the true autocorrelations are ‘negligible’ for $s' > s$, where s' is a lag larger than s . As the previous literature had regarded this variance to be just $1/N$, he stressed that ‘the effective number of degrees of freedom has been reduced’ by the factor

$$\beta = \frac{1}{\sum \lambda_{s'}^2}. \quad (5.16)$$

Narrowly, then, the EDF was defined only for the sample variance of the ACF when subsequent lags of the true ACF are small. However, this expression has apparently been taken as the inspiration for a general definition of EDF for time series. To evaluate this EDF

the infinite sum must of course be truncated (Anderson, 1983). As the ACF is symmetric about zero, has $\lambda_0 = 1$, and cannot be estimated for any value $s > (N - 1)$, the most direct expression would be

$$\beta = \left(1 + 2 \sum_{s=1}^{N-1} \lambda_s^2 \right)^{-1}. \quad (5.17)$$

where $\tau = \sum_{s=1}^{N-1} (\lambda_s)^2$ is the ‘correlation length’ (Straatsma *et al.*, 1986). However, it may be undesirable to evaluate this summation all the way to $N - 1$. Firstly, the larger the lag s , the less data are available to estimate λ_s : while $\hat{\lambda}_1$ is based on the average of $N - 1$ terms, a lag of $s = N - 1$ is based on the single product $x_1 x_N$. Also, if one truly believes the ACF falls to zero above some lag, the summation should reflect that. We revisit these issues below in section 5.1.7. For a single time series X , the EDF is obtained as

$$\hat{N}^x \approx \frac{N}{1 + 2 \sum_{s=1}^{N-1} \hat{\lambda}_s^2}. \quad (5.18)$$

We now return to the matter at hand, which is estimating the variability of Pearson’s correlation coefficient between two time series X and Y . While we do not have a theoretical justification, the conventional usage in neuroimaging suggests naive use of this EDF, assuming a common ACF for X and Y . In practice, this correction factor β could be estimated for each and then averaged, or \hat{N}^x and \hat{N}^y could be computed and averaged; or, when considering the ROIs over the brain, a \hat{N}_{Global} could be defined as the median of \hat{N} .

5.1.7 Family of Heterogeneous Bivariate Corrections (HetBiv)

As best as we can judge from the neuroimaging literature, the typical approach is to use a global EDF based on Eq. 5.18 to estimate the variance of the correlation coefficients (Fox *et al.*, 2005; Van Dijk *et al.*, 2010). This of course then fails to account for the spatial heterogeneity in autocorrelation between different nodes. Here, we propose a method of estimating an unbiased correlation coefficient of a pair of time series which allows us to account for the heterogeneity in effect of autocorrelation. Unlike the AR(1) models described in section 5.1.5, the proposed method also considers more complex patterns of autocorrelation in BOLD time series.

We first consider the case when X and Y are mean zero and variance one, such that a correlation can be estimated

$$r_{xy}^{01} = X'Y/N, \quad (5.19)$$

where we are treating X and Y as column vectors of length N . We allow arbitrary dependence, initially, such that the $N \times N$ variance-covariance matrices are $\text{var}(X) = \Sigma_x$,

$\text{var}(Y) = \Sigma_y$ and $\text{cov}(X, Y) = \Sigma_{xy}$. In this setting, we use basic results² for random vectors (Brown and Rutemiller, 1977) to find the mean

$$E(r_{xy}^{01}) = \text{Tr}(\Sigma_{xy})/N, \quad (5.20)$$

and, if we further assume multivariate normality,

$$\text{var}(r_{xy}^{01}) = \frac{1}{N^2} \left(\text{Tr}(\Sigma_x \Sigma_y) + \text{Tr}(\Sigma_{xy}^2) \right). \quad (5.21)$$

Finally, we consider real world time series that are not mean zero or variance one. As long as we confine ourselves to the case of constant variance (i.e. constant diagonals in Σ_x and Σ_y), we can then argue that these results for r_{xy}^{01} are applicable to r_{xy} (as is typically done in the time series literature, e.g. Bartlett (1946)). We are essentially assuming that, for sufficiently large N , the sampling variability in the sample means and standard deviations of X and Y are negligible relative to those of the numerator of the correlation coefficient.

Therefore, the effective degree of freedom can be obtained as

$$\hat{N}_{HetBiv-unbiased}^{xy} = N \times \left(\text{Tr}(\Sigma_x \Sigma_y)/N + \text{Tr}(\Sigma_{xy}^2)/N \right)^{-1}, \quad (5.22)$$

where we have arranged the expression to show the correction factor to be the inverse of a sum of two terms, each of which is the average of the diagonal of a matrix.

There are two special cases to consider. When X and Y are independent, Σ_{xy} is the zero matrix and we have

$$\text{var}(r_{xy}) = \frac{1}{N^2} \text{Tr}(\Sigma_x \Sigma_y). \quad (5.23)$$

Consequently, the effective degree of freedom can be determined as:

$$\hat{N}_{HetBiv-Unrelated}^{xy} = N \times \left(\text{Tr}(\Sigma_x \Sigma_y)/N \right)^{-1}. \quad (5.24)$$

Second, when X and Y are dependent but this dependence is only present in their zero-lag covariance. In this case, $\Sigma_{xy} = \rho_{xy}$, and then

$$\text{var}(r_{xy}) = \frac{1}{N^2} \left(\text{Tr}(\Sigma_x \Sigma_y) + \rho_{xy} \right). \quad (5.25)$$

Eventually, the degree of freedom can be obtained as:

$$\hat{N}_{HetBiv-0lag}^{xy} = N \times \left(\text{Tr}(\Sigma_x \Sigma_y)/N + \rho_{xy}/N \right)^{-1}. \quad (5.26)$$

²For a thorough derivation of moments of two random variables, see Appendix C.1

Calculating the methods that we propose in Eq. 5.22, Eq. 5.24 and Eq. 5.26 can be challenging. Firstly, we specify Σ_x and Σ_y as Toeplitz matrices defined by ACF's; thus, the first row of Σ_x is $(1, \lambda_1^x, \dots, \lambda_{N-1}^x)$, the second row is $(\lambda_1^x, 1, \lambda_1^x, \dots, \lambda_{N-2}^x)$. But even exploiting the Toeplitz structure, for long BOLD time series and a large number of parcellations, will be time-consuming, since the complexity of calculating HetBiv-corrections for one pair of BOLD time series for one subject is $O(2n^3)$. However, inspection of the product $\Sigma_x \Sigma_y$ reveals that the diagonal (needed for the trace) takes the form of inner products between a corresponding row and column of the two matrices. For the first element this summation is $1 + \sum_{s=1}^{N-1} \lambda_s^x \lambda_s^y$; for the middle row $(N+1)/2$ (assuming N odd), this is $1 + 2 \sum_{s=1}^{(N-1)/2} \lambda_s^x \lambda_s^y$, etc. Therefore, inspired by Bartlett's estimation method for t_s , we take Eq. 5.23 and replace $\text{Tr}(\Sigma_x \Sigma_y)$ with N times an approximation of one element of the diagonal, leading to a sample variance estimation of the bivariate correlation coefficient as:

$$\text{var}(r_{xy}) \approx \frac{1}{N} \left(1 + 2 \sum_{s=1}^{N-1} \hat{\lambda}_s^x \hat{\lambda}_s^y \right). \quad (5.27)$$

The effective degree of freedom can be estimated by using Eq. 5.27 to correct the number of observations as:

$$\hat{N}_{\text{HetBiv}}^{xy} = \frac{N}{\left(1 + 2 \sum_{s=1}^{N-1} \hat{\lambda}_s^x \hat{\lambda}_s^y \right)}. \quad (5.28)$$

We note that we have neglected the term dependent on the true cross correlations (compare Eq. 5.22 and Eq. 5.28). To ensure the impact would not bias the analysis, we conducted an empirical evaluation and found it made little impact in estimation of variance of correlation coefficients (details in Appendix C.2).

As discussed earlier, in section 5.1.6, using Eq. 5.28 in practice is not free of pitfalls. One of the problems is choosing an optimal number of lags to be involved in the estimation, as we would likely believe that after some lag the autocorrelations are truly zero and that involving all $N-1$ lags may make the estimation less-accurate due to a biased approximation of distant lags (Anderson, 1983). Therefore, the number of lags that is required to calculate the correction factor must be chosen carefully. The optimal number of lags was suggested as $N/4$ (Anderson, 1983), although we employed two separate techniques to examine how choosing the number of lags may influence the estimations.

Curbing. Instead of using all lags of the autocorrelation function, we only use an initial fraction of the lags to calculate the *HetBiv* correction factor. We examine a range of curbing coefficients which start from 10% to 80 % with increments of 10%, i.e. for the curbing coefficient of 10% we only consider the first $0.1 \times N$ ACF lags, and all the remaining lags are set to zero.

Shrinking. At each lag, we calculate the geometric mean of un-signed autocorre-

lation coefficients, $\sqrt{|\lambda_i^x||\lambda_i^y|}$, and if it exceeds a threshold we preserve the lag's autocorrelation coefficients, otherwise we set it to zero. The threshold is set relative to the nominal standard deviation of the sample autocorrelation coefficient, i.e. $1/\sqrt{N}$. We examine the thresholds for a range of multipliers from 0.25 to 3 with increment 0.25. The range was chosen as the conventional standard deviation multiplier for such studies are set to 3.

5.1.8 Monte-Carlo Correction

Beside theoretical approaches to account for the effect of the autocorrelation bias, Monte-Carlo methods were also shown to be useful (Kalos and Whitlock, 2008). FSLnets³ is one of the most widely-used toolboxes which estimates the standard error of sample correlation coefficients by employing a Monte-Carlo simulation method. For each node, the AR(1) coefficient of the BOLD time series is computed by Eq. 5.8 and then the median of all of these Φ_1 values is used to simulate data from an AR(1) process, X^{MC} , using Eq. 5.7. The Pearson's correlation coefficients between each pair of surrogate timeseries, r^{MC} , is calculated and then transformed to Fisher's \hat{z}^{MC} . For a large sample of \hat{Z}^{MC} , the variance is computed and used to derive a global correction for all of the time series for that subject. It is also worth noting that FSLnets only generate as many surrogate time series as the number of original time series, and thus may give a poor estimate of the standard deviation if a small number of nodes is considered.

5.1.9 Method Comparison

The conventional approach of evaluating EDF is to use the simulation of a simple model (normally first order) autoregressive function, but it is known that BOLD signals are of a more complicated temporal structure (Logothetis and Wandell, 2004), which makes it essential to avoid using simplistic temporal models for such an evaluation. Therefore, in this study we use BOLD signals to evaluate the proposed methods instead of simulating synthetic time series aiming to ensure that our methods are tested under close-to-reality complications of a time series structure.

Due to the non-reproducible nature of BOLD signals, the correlation coefficient between a time series from subject k and a time series from subject q is supposed to be zero (Smith *et al.*, 2011b). Following this realisation, by repeating the process of correlating randomly selected nodes from one subject to a randomly selected node from another subject we expect to form a mean-zero univariate normal distribution unless a non-random systematic confound, such as autocorrelation, inflates correlation coefficients. The correlation coefficient between two time series with systematic non-white noise results in a non-zero

³<http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSLNets>; visited on 18 September 2015

correlation coefficient due to non-physiological grounds which were mentioned in the introduction of this chapter.

To examine, the difference between the shape of distributions, we use two-sided unequal Kolmogorov-Smirnov statistics (Daniel *et al.*, 1990), K which is defined as the maximum of the absolute difference between the Cumulative Distribution Function (CDF) of the surrogate data, Ψ , and a mean-zero univariate normal distribution, \mathcal{N} , and can be formulated as below:

$$K = \max_x (|\Psi(x) - \mathcal{N}(x)|). \quad (5.29)$$

Hence, a value of K close to zero indicates the shorter the distance from a normal distribution.

In addition to examining the shape, it is also crucial to ensure that the distribution tails do not corrupt the statistical tests. Therefore, we examine how well each of these methods can transform the distribution tails. We approximate the probability of inter-subject inter-node correlation coefficients being bigger than a z-score critical value of 1.96 for a two-sided p-value of 2.5%, and 2.57 for two-sided p-value of 0.5%. Methods with higher probability of producing values larger than 1.96 and 2.57 are considered to be ‘conservative’ methods and those with lower probability of producing values lower than 1.96 and 2.57 are considered as ‘liberal’ methods.

5.1.10 Network Thresholding

Statistical Thresholding

In statistical thresholding (ST) we test each correlation coefficient against the null hypothesis that the observed correlation coefficient is equal to zero. We use Fisher’s transformation \hat{z} to ensure normality in the correlation coefficient. To perform the test, we use standard scores, \bar{z} , which are obtained from multiplying \hat{z} by the standard error obtained from an appropriate effective degree of freedom \hat{N} . P-values of each edge, representing the probability of the correlation coefficient to be zero, are obtained by checking the standard score against a univariate normal distribution, \mathcal{N} . To correct for false positives due to Multiple Comparison Error (MCE), P-values of edges with non-zero correlation coefficients ($\leq \frac{K(K-1)}{2}$) undergo False Discovery Rate (FDR) correction. FDR is a correction method which dynamically adapts the rejection threshold according to the p-values. Finally, to form binary networks, any correlation coefficients that survives the hypothesis test and the MCE correction, is set to one and any of them that fail to reject the hypothesis are set to zero.

Density Thresholding

Density thresholding (DT) is a widely used method for thresholding adjacency matrices over a range of densities. In DT, edge weights of an adjacency matrix are thresholded by a certain value, therefore, supra-threshold edges are preserved while the rest are set to zero. In this setting, only strongly connected edges are preserved. Since the true density of the functional connectome of the human brain is unknown, we study the graph phenotypes over a range of thresholds. The density range starts from a large threshold which only preserves a set of highly connected edges and proceeds to a modest thresholds which preserve a larger number of supra-threshold edges. It is worth noting that one should be careful when selecting the boundaries as a very small lower-boundary results in a large number of disconnected components or isolated nodes, while a liberal upper-boundaries leads the graph to become a lattice. We use eight consecutive density steps starting from a density of 5% to 45% with an incremental step of 5%, the reason behind the choice of density choice was discussed in details in Chapter 3. In each density, the correlation coefficients are transformed to Fisher values and then standardised by an appropriate degree of freedom.

5.1.11 Graph Theoretical Measures

To investigate how inflated degree of freedom, due to non-white noise, can affect the description of the functional human connectome, we utilise a range of traditional graph theoretical matrices, which are commonly used to describe topological properties of the human brain. In this chapter, we use clustering coefficients and normalised clustering coefficients as measures of segregation (see section 3.1.4), characteristic path length and normalised characteristic path length as measures of integration (see section 3.1.3) and small-worldness as a measure of complexity (see section 3.1.6).

5.2 Data

In this section, we briefly describe the properties of each cohort that we used throughout this study including the acquisition procedures and the sample demographics for each subject.

5.2.1 Human Connectome Project

Resting-state fMRI scans of 318 subjects (mean age=29, \pm 7 years; female=132) were obtained through the Human Connectome Project consortium. The cohort has 78 monozygotic twins and 69 dizygotic twin pairs. Participants were asked to remain awake in a MRI machine with their eyes open. Each subject was asked to fixate on a projected bright cross-hair on a dark background. Gradient-echo EPI (TE=33.1ms, flip angle=52 °, FOV=208x180mm)

was used to acquire a 15min scan of each subject with $TR=720ms$, which result in time series of 1200 data-points for each subject. Each brain volume was parcellated using isomorphic voxels of $2mm^3$.

We used an extended pre-processed version of the data available via the HCP repository. Each scan has undergone a thorough pre-processing procedure as explained in [Glasser *et al.* \(2013\)](#)⁴. Further to the traditional pre-processing steps, high-pass filtered data were denoised by the non-aggressive FIX algorithm ([Salimi-Khorshidi *et al.*, 2014](#)) to ensure that any 'bad' components, such as movement, physiological noises of a heart beat and respiration, have been regressed out of each individual scan. Since there is an ongoing debate about the effect of global signal regression (GSR), we also repeated the analysis with and without global signal regression.

5.2.2 Autism Brain Imaging Data Exchange

For replicating the analysis we used the resting-state fMRI of healthy subjects, publicly available, through Autism Brain Imaging Data Exchange (ABIDE) ([Di Martino *et al.*, 2014](#)). It is important to note that despite the name of the cohort (Autism) we only use the healthy subjects. Also, the cohorts were chosen in a way that they cover a different range of sample size and scan length as it is crucial to examine robustness of our methods under different circumstances. The data are from three separate cohorts:

- New York University Lagone Medical Centre (NYU): 86 healthy subjects (mean age= 15 ± 4.5 years; female=32) were asked to remain awake with their eyes open. Resting-state fMRI data were acquired for length of 6.40 min, with $TE=25ms$ and $TR=2s$, Flip Angle: 60° , voxel size= $3.4 \times 3.4 \times 4mm$ and Anterior to Posterior k-space phase encoding. Therefore, the BOLD signal of each voxel was represented by a time series of 175 data-points ([Chabernaud *et al.*, 2012](#)).
- Utah School of Medicine (USM): 23 healthy subjects (mean age= 21 ± 12 years; female=9) were asked to remain awake with their eyes open. Resting-state fMRI data were acquired for length of 6 minutes, with $TE=15ms$ and $TR=2s$, Flip Angle: 90° , voxel size= $3 \times 3 \times 4mm$ and Right to Left k-space phase encoding. Therefore, the BOLD signal of each voxel was represented by a time series of 235 data-points ([Anderson *et al.*, 2010](#)).
- Yale School of Medicine (Yale): 25 healthy subjects (mean age= 13 ± 5 years; female=10) were asked to remain awake with their eyes open. Resting-state fMRI data were acquired for length of 8.06 minutes, with $TE=28ms$ and $TR=2s$, Flip Angle: 90° .

⁴For general description of pre-processing stages required in analysis of fMRI data see section 2.1

°, voxel size= $3.4 \times 3.4 \times 4\text{mm}$ and Right to Left k-space phase encoding. Therefore, the BOLD signal of each voxel was represented by a time series of 195 data-points (Di Martino *et al.*, 2014).

We used the pre-processed version of each cohort which is publicly available through the Pre-processed Connectome Project (PCP) ⁵. Pre-processing was done by the Connectome Computation System (CCS) pipeline (Xu *et al.*, 2015). The first four volumes were discarded as the machine is reaching the equilibrium. Slice timing correction and motion realignment were performed as basic processing measures. In addition, 24 motion parameters, mean White Matter and Cerebrospinal Fluid (CSF) were regressed out. Eventually, a band-pass filter (0.01 – 0.1Hz) was used to remove the high and low power noise.

5.2.3 Brain Parcellation Schemes

In order to investigate the relationships between different parts of the brain, we use four different parcellation schemes to decimate a voxel-wise volume of the brain into a collection of Regions of Interest (ROI). Parcellation schemes which were used in this study are from two major categories of parcellation schemes, data-driven and anatomical. Anatomical atlases subdivide the brain according to previously known features of the human brain anatomy (Desikan *et al.*, 2006; Eickhoff *et al.*, 2007), whereas, data-driven atlases parcellated the brain based on the similarity in pattern of the functional activities among the voxels (Craddock *et al.*, 2013). In the following we discussed each parcellation scheme briefly:

- **Yeo17** is a data-driven atlas which was formed out of similar patterns of functional connectivity of voxels across two resting-state scans of 500-subject cohorts of healthy brains. Networks of functionally coupled regions were identified by employing a machine-learning clustering method (Yeo *et al.*, 2011). Resultant ROIs are dropped into 7 Resting-state networks (RSN): Default Mode Network, Dorsal Attention Network, Saliency Ventral Attention Network, Control Network, Limbic regions, Visual and Somatomotor Networks. However, these 7 RSNs, are also subdivided into 17 networks. These 17 sub-networks were found across a surface-based realignment of all subjects which merely allows investigation of the cortical regions of the human brain. This parcellation scheme delivers clear evidence of a segregation/integration nature of the human brain function as well as hierarchical structures of the connectivity profiles.
- **ICA100** is a data-driven atlas which was formed by applying Independent Component Analysis (ICA) on resting-state functional scans of the human brain. ICA is a

⁵<http://preprocessed-connectomes-project.org/> ; visited on 28 July 2016

mathematical method of distinguishing between the latent number of additive components of a multivariate signal (Beckmann *et al.*, 2009). Applying ICA on the brain fMRI results in spatially segregated regions of the brain which are involved in a specific function. This leads the components to be subjects to a level of modularity in advance. Since the true number of components is unknown, choosing the number of components allows a trade-off between size and the number of ROIs. For this study, we used time series which were already extracted and are publicly available through HCP consortium. For extracting this time series, between a Group-PCA of 820 subjects was generated by MELODIC's incremental Group-PCA and then fed into a group-ICA using FSL's MELODIC tool for a dimensionality of 100 (Smith *et al.*, 2013). It is worth noting that not all ICA components are necessarily physiologically meaningful, some of them may have been caused by head movement or cardiac processes, but considering that the 'bad' components were already extracted from the data by FIX-ICA (Salimi-Khorshidi *et al.*, 2014), all the 100 components are expected to be carrying a large range of physiological information. ICA-100 covers the cortical, sub cortical and cerebellum regions of the human brain. As the RSN assignment of nodes is unknown, we discard local examination of changes for this specific parcellation scheme.

- **Power2011** is a data-driven atlas which is formed by same-sized spheres around the putative functional coordinates (Power *et al.*, 2011). These putative coordinates are a combination of task-related and resting-state regions of the human brain. Through meta-analysis of more than 72 studies, 152 functional putative coordinates related to task fMRI were found. Furthermore, 193 coordinates were also found by analysing a healthy resting-state cohort of 1000 subjects. By removing the overlaps between the coordinates found by these two methods, 264 coordinates remain as representatives of the functional activities within the cortical, sub cortical and cerebellum of the human brain. This brain parcellation has relatively small and same-sized regions of interest which helps avoid the possible overlaps between signals of different regions. Regions of interest in Power2011 are divided into 12 RSN networks including the Sensory, Cingulo-opercular task control, Auditor, Default Mode Network, Ventral Attention Network, Visual, Fronto-parietal task control, Salience network, subcortical network, Cerebellar network and Dorsal Attention Network.
- **AAL** is an anatomical parcellation scheme which is based on a high-resolution T1-weighted MRI scan of a single-subject. In the initial version of the AAL, each hemisphere was divided into 45 ROIs. Each of these ROIs, is the result of manual 2D drawing of boundaries between the main sulci (Tzourio-Mazoyer *et al.*, 2002). The

2D boundaries of each slice then formed a 3D volumetric ROI. However, since its inception, it has undergone two major improvements: [1] 26 new ROIs were added to the original 90-ROI atlas (Tzourio-Mazoyer *et al.*, 2002); [2] 4 additional new ROIs were added in the Fronto-Parietal lobe (Rolls *et al.*, 2015). Here, we use the latest version, conventionally called AA2, with 120 ROIs. This parcellation scheme covers the cortical, subcortical and cerebellum of the human brain. Although the AAL2 is an anatomical atlas, the RSN assignments of each node are available via cross-matching of the ROIs with other data-driven parcellation schemes. The AAL2 resting-state networks include 8 RSNs: The Default Mode Network, Visual and Visual Lateral, Fronto-parietal, Auditory, Sensorimotor, Executive Control and Cerebellum.

- **CC200** is a data-driven atlas which is formed by 200 ROIs of similar sizes. Using forty-one healthy subjects, individual functional connectivity was formed for each subject by measuring the temporal similarity between each pair of voxels. Supra-threshold edges were preserved while the rest of the connectivity map was set to zero. Normalised cut spectral clustering was used to parcellate each subject's connectivity map. Then for each subject, a co-occurrence matrix was formed indicating which nodes are partitioned together. Finally, a parcellation scheme were formed by partitioning the group-averaged co-occurrence matrices into 200 modules (Craddock *et al.*, 2012). In this work, we use CC200 only to examine the correlation lengths and validation processes.

5.3 Results

5.3.1 Effect of Autocorrelation on Brain Cortical and Sub-cortical Regions

To examine the effect of autocorrelation on the cortical and subcortical regions of the brain we calculate the correlation length of subjects from the HCP cohort. Figure 5.1 shows the average correlation length across the subjects. It suggests that the effect of autocorrelation is not homogeneously distributed over the brain. For example, the effect of autocorrelation in the Parietal and the Occipital lobes is relatively larger than the Frontal lobe. This effect is almost negligible in the Temporal lobe and the Cerebellum and Limbic areas.

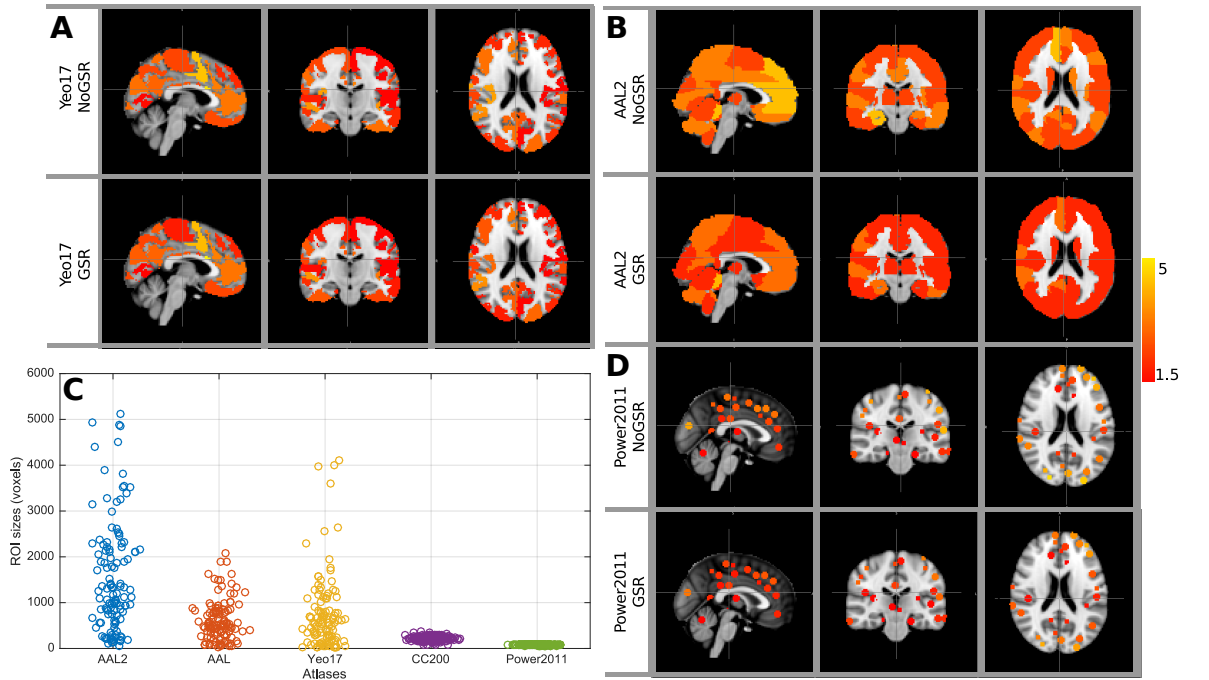


Figure 5.1: A,B,D Group-averaged correlation length for Yeo17, AAL2 and Power2011, respectively. C. Volume of the ROIs across the parcellation schemes.

Further to measuring the correlation length of each parcellation scheme, we also examined to what extent they are associated with the size of each ROI. For each parcellation scheme, we calculated the correlation length of each ROI across subjects and correlated it with the ROI sizes. Figure 5.2 shows the level of association between the ROI sizes and the effect of autocorrelation. Results suggest that a high positive correlation between the size of the ROIs and correlation length in the parcellation schemes with relatively large ROIs (please note the results for Yeo17, AAL and AAL2). However, this association fades for parcellation schemes with a higher number of ROIs and a smaller volume of ROIs (please

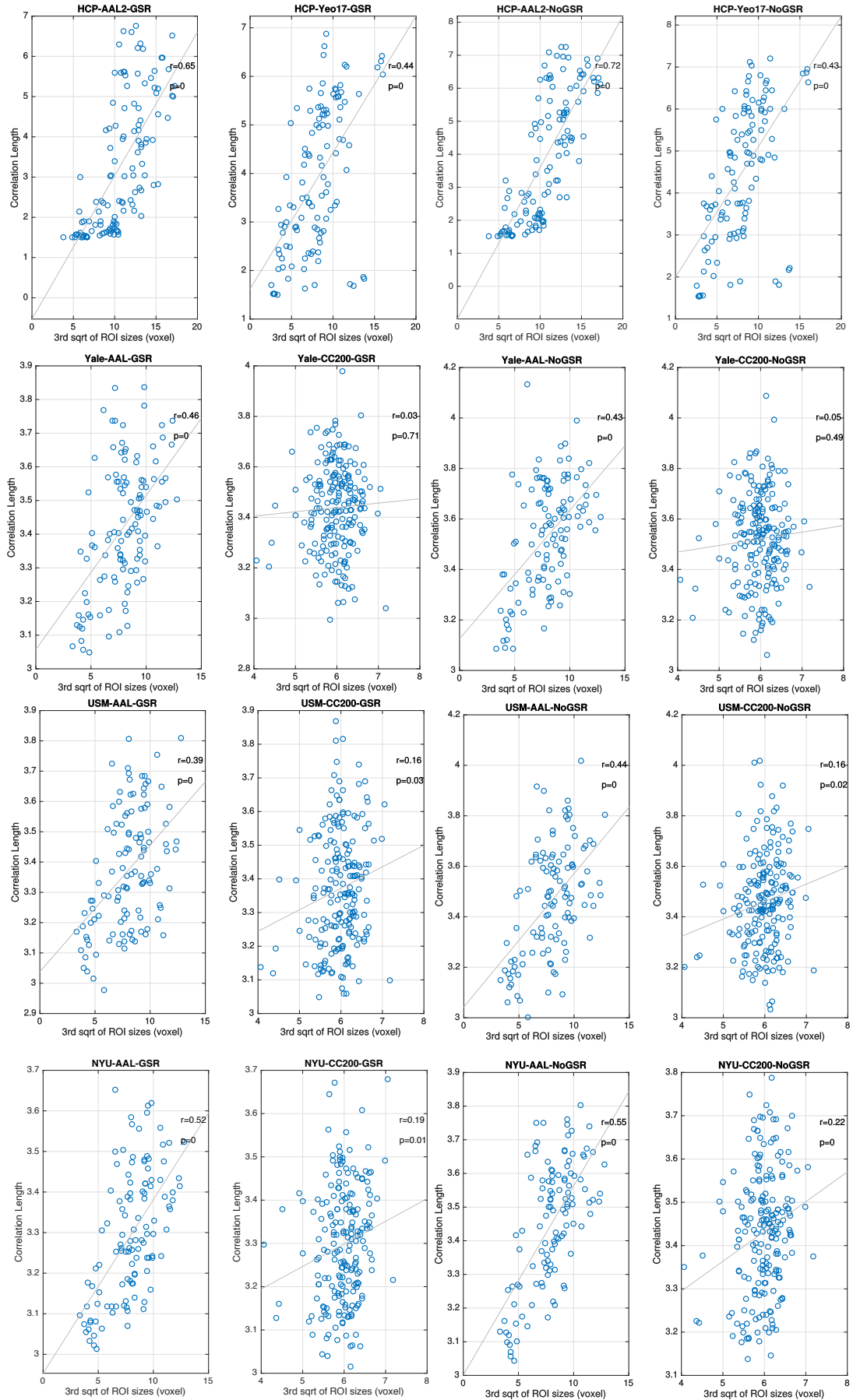


Figure 5.2: Association between correlation length and four different parcellation schemes (e.g. Yeo17, AAL2, AAL, CC200) of three different scanning sites (e.g. Yale, NYU, USM and HCP) with and without GSR.

note the results for CC200).

5.3.2 Results of the Correction Methods Validation

In this section, we show how different methods of estimating the degree of freedom performs on real-data validations. We separately compare methods for their distribution shape and false-positive rate (tail analysis).

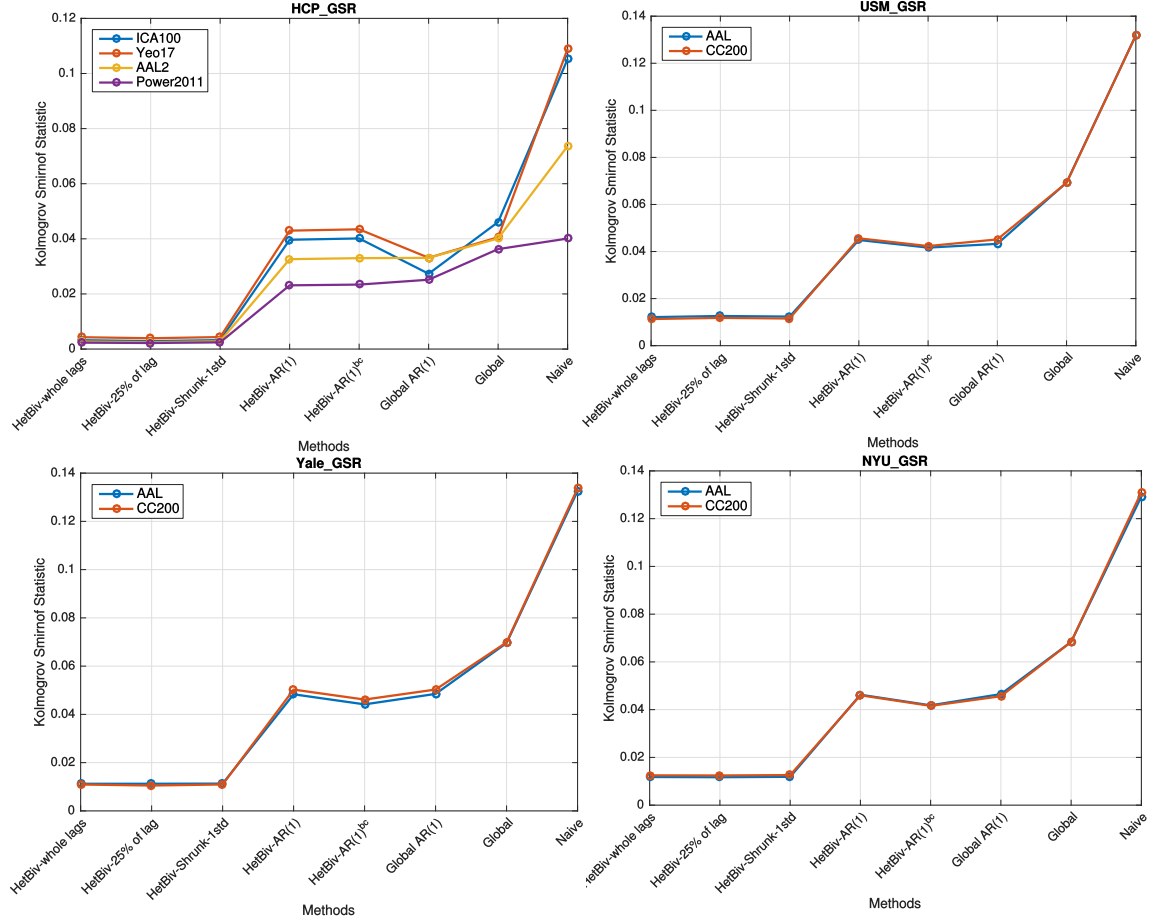


Figure 5.3: KS statistics of a family of HetBiv corrections as well as existing methods for each acquisition site and parcellation scheme. The grey lines represent the reference line.

Figure 5.3 shows the KS statistics between the estimated distribution and the normal distribution for each method. In the HCP site, with a long time series, the class of HetBiv corrections outperforms existing local and global methods. Among the HetBiv methods, HetBiv-Curbed (25%) and HetBiv-Shrunk (1σ) show similar results to HetBiv methods without curbing and shrinking. In the HCP cohort, the family of HetBiv corrections shows similar results across parcellation schemes, whereas, for the other methods of

correction, each parcellation scheme has a different KS statistics. For instance, Power2011 has a relatively lower KS statistic, whereas, Yeo17 has relatively a higher KS statistic for the non-HetBiv correction methods.

In the NYU, USM and Yale cohorts, with shorter time series, the HetBiv corrections still appears to be the best method of estimating the effective degree of freedom. It is notable that, contrary to the HCP cohorts, parcellation schemes in the three cohorts demonstrate a relatively similar KS statistics.

We also examined the effect of different curbing coefficients (Appendix C.3) and shrinking factors (Appendix C.4) across a wide range of values. The results suggest that although curbing and shirking can, in some cases, improve the performance of the HetBiv, it is not consistent across parcellation schemes and cohorts.

In addition to examining the distribution shape, we also investigated the tails of the distributions aiming to ensure that the correction methods have properly corrected the tails as it is crucial in further statistical inferences required in statistical thresholding. Figure 5.4 shows how each correction method performs in correcting the distribution tails. The red and blue lines represent α -levels of 0.05 and 0.01, respectively. Any method that exhibits an α -level beyond these lines is considered 'conservative' since it will lead the test to produce a level of false-negative and any method with an α -level lower than the blue and the red lines is considered to be 'liberal' as it will lead to a level of false-positive.

For the HCP site, the naive correction method results in the highest level of false-positive. It is followed by the global AR(1), HetBiv AR(1) and HetBiv AR(1)-BC. Contrary to these methods, the Global HetBiv is the most conservative method which may cause false-negatives in the analysis. Three remaining methods, HetBiv-all lags, HetBiv shrunk and HetBiv curbed are located on the border lines which implies that these three methods produce the results with the highest level of sensitivity and specificity. For the three remaining sites, NYU USM and Yale, the naive correction is most likely to produce false-positives. Two other global methods, Global HetBiv and Global AR(1) as well as local methods, are most likely to produce false-negative results due to the conservatism. The remaining HetBiv methods, are most likely to have the closest-to-truth sensitivity and specificity. It is also worth noting that the naive method for shorter time series (scans from NYU, USM and Yale) are more likely to result in false positives. The parcellation schemes that enjoy a large number of the ROIs and higher volume similarity across ROIs have a relatively lower rate of false positive (see results of CC200 and Power2011 in Figure 5.4).

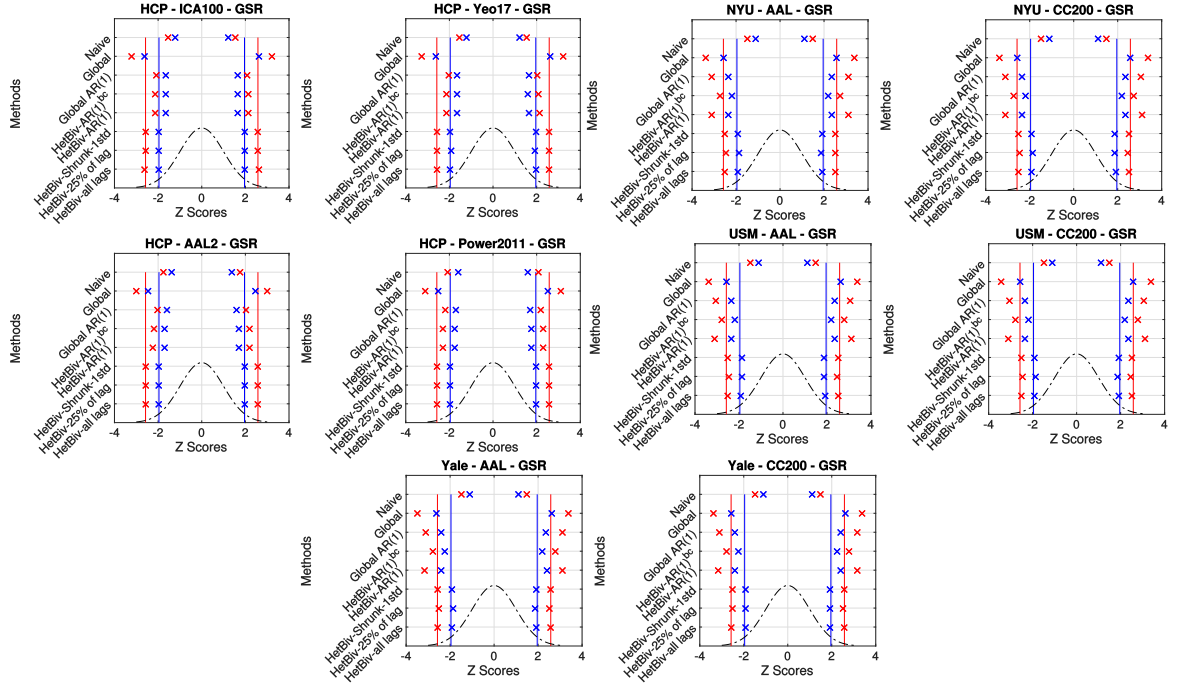


Figure 5.4: False-positive rate of family of HetBiv corrections as well as existing methods for each acquisition site and parcellation scheme. The acquisition site has been identified by the titles.

5.3.3 Effect of Autocorrelation Correction on Statistical-Thresholding

To examine the effect of correcting the degree of freedom on statistical thresholding, we only use the Resting-state fMRI scans from the HCP cohort to investigate how the correction can affect the correlation coefficients between nodes. Figure 5.5 shows \hat{z} values of four parcellation schemes with Naive correction and HetBiv-correction. The figure also shows the effect of GSR across the parcellation schemes. Results suggest that, firstly, the GSR can drastically transform the \hat{z} values. Before regressing out the mean value, the \hat{z} scores are mostly distributed on the positive side of the histogram, however, after GSR, the distributions are centred around zero which results in the larger number of anti-correlations. It is notable that the histograms of \hat{z} scores for ICA100 suggest no differences with and without GSR. Figure 5.5 also suggests that the HetBiv-correction can also affect the \hat{z} values, consistently for all parcellations, as it eliminates the weak correlation coefficients and enlarges the \hat{z} values of the larger correlation coefficients.

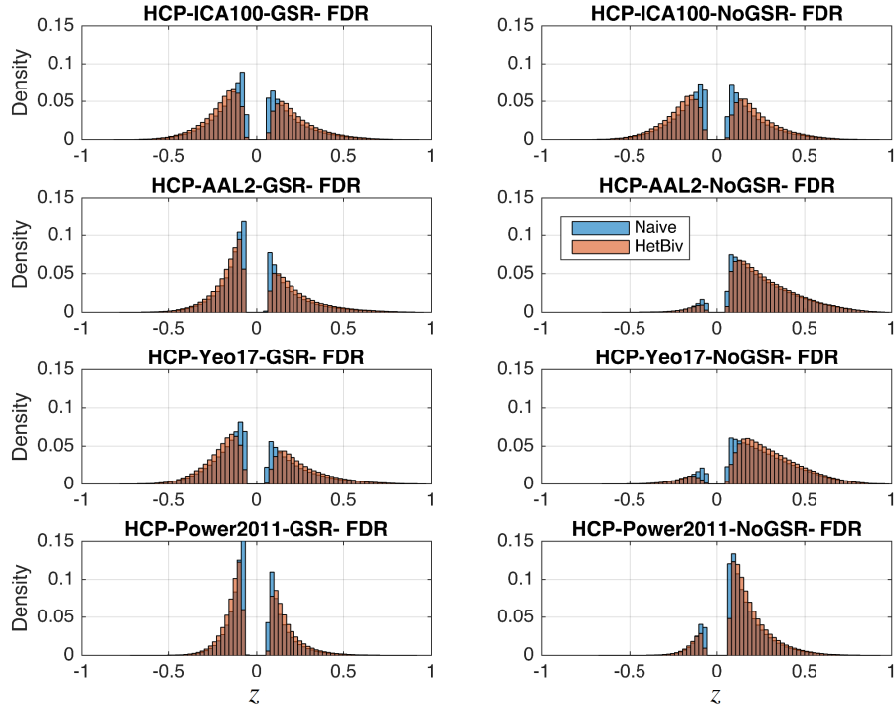


Figure 5.5: Histogram of the Fisher-transformed correlation coefficients after density thresholding with Naive and HetBiv correction for the degree of freedom.

Correcting the degree of freedom can have an immediate effect in the global density of networks after statistical thresholding. A smaller degree of freedom results in a more conservative test which can lead the network to have a lower number of supra-threshold edges. To examine the effect of the correcting of DoF on the network density, we also calculated the density of each resultant network for Naive-correction and HetBiv-correction. Additionally, we also examined how GSR can affect the network density. The result of the network density is shown in Figure 5.6. The HetBiv-correction results in a more conservative test and, consequently, a lower density in statistical thresholding. It is notable that the GSR not only decreases the networks density, it also reduces the variation between subjects.

5.3.4 Effect of Autocorrelation Correction on Graph Theoretical Measures of Statistical-Thresholding

In this section, we examine the effect of Naive and HetBiv techniques to estimate the effective degree of freedom in networks with and without global signal regression. We examine this effect on global metrics (i.e. segregation, integration and complexity) along with local

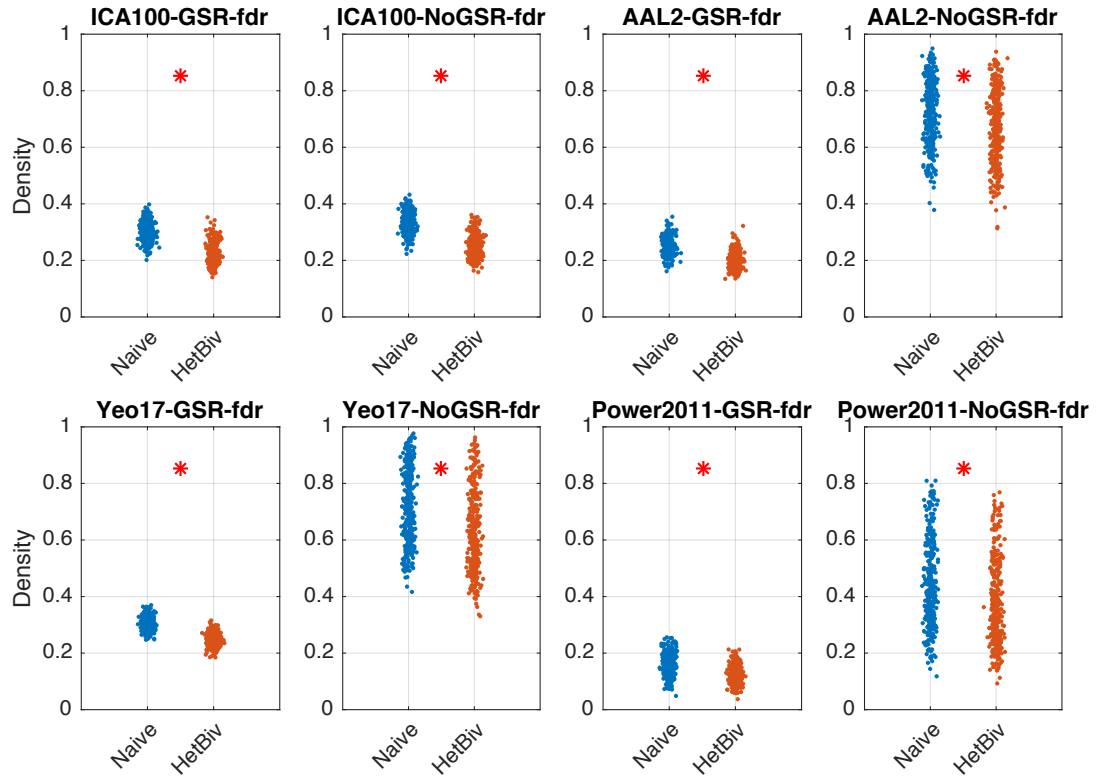


Figure 5.6: Network density of networks with Naive and HetBiv corrections. Each subplot shows the results for the choice of GSR and parcellation scheme. The red asterisk shows statistically significant results.

metrics (i.e. centrality measures and Rich Club coefficients).

Global Graph Theoretical Measures

The network segregation level is shown in Figure 5.8 by the mean clustering coefficient (denoted by MC in Figure 5.8), normalised mean clustering coefficient (denoted by letter O in Figure 5.8) and the modularity index (denoted by MI in Figure 5.8). The results suggest a significant decrease in the average clustering coefficient after the HetBiv correction. However, the normalised clustering coefficients show a significant increase due to HetBiv-correction. Another measure of segregation, the modularity index, is also increased due to the HetBiv correction. It is also notable that global signal regression increases the level of segregation in networks for both cases of Naive and HetBiv corrections.

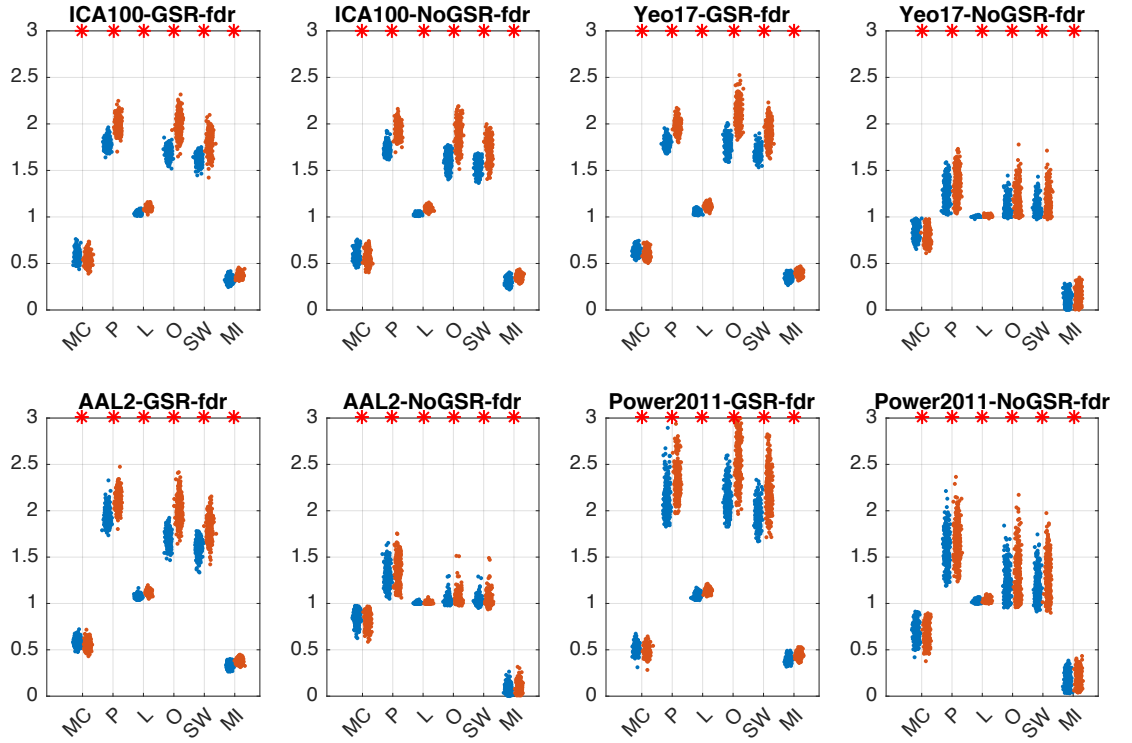


Figure 5.7: Global graph theoretical measures of networks with Naive and HetBiv corrections. Each subplot shows the results for choice of GSR and parcellation scheme. Red asterisk shows statistically significant results.

The level of network integration is measured by the characteristic path length (P) and it is a normalised version (Λ ; L). Figure 5.8 shows that the non-normalised characteristic path length is significantly increased due to the HetBiv-correction, however this difference is reduced later in normalised values. The results suggest that, despite the small differences, traversing in a functional network will be more costly with HetBiv-correction. The figure also suggests that global signal regression may deflect the level of network integration by increasing the normalised characteristic path length.

The complexity level, which is measured by a small-world index (SW), shown in Figure 5.8, suggests that the level of complexity in a network increases after using the HetBiv-correction. The increase in SW has been led by a drastic increase in the level of segregation while the level of integration approximately stays as high as a random graph. It is also notable that the global signal regression increases the level of complexity in a network across all four parcellation schemes.

Local Graph Theoretical Measures

For quantifying the effect the HetBiv-correction on a local level, we use centrality and coreness measures such as Degree Centrality (DC), Betweenness Centrality (BC) and non-normalised/normalised Rich Club coefficients (RC and NRC). We use Disruption Plots to summarise the local changes due to Naive and HetBiv corrections ⁶. We discard ICA100 from further local analysis as the RSN assignment of nodes in this parcellation scheme is unknown. The disruption indices are summarised in Table 5.1 and Table 5.2 for the case of the GSR and NoGSR, respectively. Each disruption index corresponds to a RSN of a relative parcellation scheme which is later identified by different colours in Figure 5.8.

GSR - Statistical Thresholding

	Degree Centrality	Betweenness Centrality	Rich Club Coefficients	Normalised Rich Club Coefficients
Power2011				
Uncertain	-0.32	0.15	-0.12	0.46
Sensory/Somato	-0.31	0.13	-0.16	0.23
Cingulo Control	-0.25	0.17	-0.05	0.21
Auditory	-0.25	0.1	-0.1	0.24
Default Mode	-0.22	0.09	-0.05	0.36
Ventral Attention	-0.31	0.12	-0.15	-0.14
Visual	-0.25	0.2	-0.07	0.25
Fronto-parietal	-0.23	0.1	-0.09	0.26
Salience	-0.28	0.14	-0.1	0.32
Subcortical	-0.27	0.3	-0.03	0
Cerebellar	-0.23	0.18	0	0.02
Dorsal Attention	-0.29	0.03	-0.15	0
AAL2				
Sensorimotor	-0.18	0.24	-0.03	0
Executive Control	-0.22	0.08	-0.01	0.07
Fronto-parietal	-0.26	0.12	-0.11	0
Auditory	-0.21	0.18	-0.03	0.15
Default Mode	-0.29	0.04	-0.17	0.25
Visual	0.22	0.04	-0.21	-0.04
Visual Lateral	-0.21	0.66	-0.19	0
Cerebellum	-0.2	0.14	-0.02	0.18
Yeo17				
Visual	0.15	-0.02	-0.06	0
Sensory	-0.27	-0.25	-0.09	0
Dorsal Attention	-0.14	0.19	0	0
Ventral Attention	-0.13	0.1	-0.03	0.09
Limbic	-0.31	0.47	-0.24	0
Control	-0.23	0.14	-0.11	0
Default Mode	-0.14	0.18	-0.06	0.32

Table 5.1: Disruption indices, as discussed in section 3.1.10, of RSNs of three atlases (Power2011, AAL2 and Yeo17) for the NoGSR case of local measures in Statistical Thresholding.

⁶For more information about Disruption Plots, see section 3.1.10

NoGSR - Statistical Thresholding

	Degree Centrality	Betweenness Centrality	Rich Club Coefficients	Normalised Rich Club Coefficients
Power2011				
Uncertain	-0.16	0.14	-0.03	0.27
Sensory/Somato	-0.15	0.28	0	0.32
Cingulo Control	-0.11	0.15	0.05	0.31
Auditory	-0.11	0.15	0.1	0.34
Default Mode	-0.16	0.14	0	0.25
Ventral Attention	-0.1	0.11	0.12	0.46
Visual	-0.12	0.13	0.13	0.36
Fronto-parietal	-0.1	0.11	0.09	0.35
Salience	-0.13	0.16	-0.02	0.39
Subcortical	-0.02	0.27	0.13	0.37
Cerebellar	-0.19	0.16	0.09	0.02
Dorsal Attention	-0.17	0.07	0.04	0.35
AAL2				
Sensorimotor	0.25	0.43	0.59	0.35
Executive Control	-0.01	0.12	0.12	0.43
Fronto-parietal	0.1	0.41	0.24	0.42
Auditory	0.04	0.27	0.24	0.2
Default Mode	0.02	0.13	0.23	0.56
Visual	0.26	0.2	0.68	0.03
Visual Lateral	0.34	-0.95	0	0.41
Cerebellum	0.09	0.27	0.13	0.37
Yeo17				
Visual	0.37	0	0.82	0.76
Sensory	0.37	0.12	0.09	0.93
Dorsal Attention	-0.05	0.18	0.24	0.59
Ventral Attention	0.06	0.31	0.25	0.67
Limbic	-0.02	0.52	0.24	0.57
Control	0.03	0.28	0.2	0.99
Default Mode	-0.01	0.38	0.24	0.61

Table 5.2: Disruption indices, as discussed in section 3.1.10, of RSNs of three atlases (Power2011, AAL2 and Yeo17) for the NoGSR case of local measures in Statistical Thresholding.

Degree Centrality

In the analysis of Degree Centrality, the slope of the fitted line to the result in Figure 5.8 suggests that the effect of correcting for degree of freedom is associated with the rate of connectivity of the nodes. As highly connected nodes, degree hubs, experience the largest disruption and nodes with relatively lower rate of connectivity experience a smaller effect. It is also worth noting that almost all of the nodes have experienced a significant decrease.

- Power2011, in case of GSR, all of the RSNs experienced a decrease. However, the rate of change is not consistent across RSNs. For instance, the Cerebral and Subcortical nodes are among those with the lowest level of decrease, whereas, the Fronto-parietal, Salience, DMN, Dorsal attention and Visual networks have experienced the largest decrease. The association between the rate of connectivity and the decrease

due to HetBiv-correction is around ≈ 0.30 . Similarly, in the case of NoGSR, the results suggest a unanimous decrease across RSNs. However, nodes of the Cerebral and Subcortical networks experienced the least decrease, whereas, the Fronto-parietal, Salience, DMN, Dorsal attention and Visual networks experienced the largest decrease. It is notable that the Visual, DMN, Salience and Sensory networks contain the degree hubs. The association between the rate of connectivity and the decrease due to HetBiv-correction, in the case of NoGSR, is ≈ 0.15 .

- Yeo17, in the case of GSR, the results suggest a decrease in all DC values of all RSNs. However, the rate of decrease is not consistent as the Saliency Ventral Attention, Dorsal Attention, DMN and Control networks suggest the highest decrease, whereas, the Limbic system experienced the lowest decrease. The hub degrees are among the DMN, Saliency and Dorsal attention networks. In the case of NoGSR, similar to the GSR case, all of the RSNs show a decrease in their DC values, however, this decrease is suggested to be lower in Limbic system and higher in the Dorsal Attention, Somato-sensory, DMN networks.
- AAL2, in the case of GSR, the DC values of the nodes within all of the RSNs experience a decrease. However, the Cerebellum and Executive Control are among the least affected, while, the Fronto-parietal, Sensorimotor and Visual Lateral are among the most affected RSNs. The association between the DC values and the effect of HetBiv-correction is ≈ 0.20 . Similarly, in the case of NoGSR, the DC values of all of the nodes experience a decrease. The Cerebellum, Visual Lateral and Auditory networks are among the RSNs which have the least effect, whereas, the Visual and Sensorimotor networks are among the most affected RSNs. In the case of NoGSR, the RSNs demonstrate a negligible level of association between the DC values and changes due to HetBiv-correction.

Betweenness Centrality

The second row of Figure 5.8 shows the changes in Betweenness Centrality of Naive-corrected networks and HetBiv-corrected networks. It suggests that the HetBiv-correction introduces a positive association between the BC values of nodes and the magnitude of the changes due to the use of HetBiv-correction. In the case of GSR, 70%, 82%, 71% of nodes in Power2011, Yeo17 and AAL2 are showing a significant difference, respectively. The percentages in the case of NoGSR are 48%, 92%, 56% respectively.

- Power2011, in the case of the GSR, apart from the Memory Retrieval and Sensorimotor networks which suggest a combined increase and decrease in the BC value of

the nodes within them, the rest of the RSNs suggest a unanimous increase in their BC values. The Ventral Attention, Dorsal Attention and Fronto-parietal and Salience networks are among the networks with highest increase, whereas, the Cerebral and Subcortical regions are among the RSNs with least increase. It is also worth noting that all RSNs suggest a positive association between BC values and the effect of HetBiv-correction, however, this effect is often negligible. In the case of the NoGSR, similar to the GSR, in addition to the Memory Retrieval and Sensorimotor, Visual network also suggest a combined increase and decrease in BC values. The remaining RSNs, consistently, suggest a increase in BC values. For instance, Ventral attention shows the highest change in BC values, whereas, the Cerebral and Subcortical regions suggest a relatively lower increase in their BC values.

- Yeo17, in the case of GSR, the Somatosensory network as well as the Visual network fail to present a consistent changes in their nodes as some of the nodes suggest an increase in BC and the remaining suggest otherwise. However, the remaining RSNs suggest an increase in BC values. It is worth noting that the direction of association is not consistent across the RSNs as, for instance, the Somatosensory network suggests a negative association, while, Limbic system suggest a strong positive association. In the case of the NoGSR, apart from the Somatosensory network which demonstrates a combined increase and decrease in BC, the remaining RSNs suggest an increase in their BC values. The largest increase in BC value of an RSN belongs to the Limbic system.
- AAL2, in the case of GSR, there are only two RSNs (Sensorimotor and Visual networks) which suggest a mixture of increase and decrease in their BC values. The remaining RSNs suggest an increase in the BC. In the case of NoGSR, it is only the Visual network which demonstrates a combined increase and decrease in BC. Otherwise, the remaining RSNs suggest a increase in their BC values.

Normalised and Non-normalised Nodal Rich Club Coefficients

The last two rows of figure 5.8 show changes in nodal RC and nodal normalised RC values across the RSNs. The majority of nodes experience a significant change in their nodal RC values as the percentage of significantly different nodes in the case of GSR are 97%, 89%, and 100% for Power2011, Yeo17 and AAL, respectively. In the case of NoGSR, these percentages are 88%, 95% and 87%, respectively. Nodal NRC values are subject to the lower proportion of significant differences since in the GSR case these percentages are as low as 40%, 15% and 13%, whereas, in the NoGSR case, these percentages are 72%, 93% and 88%, respectively.

- Power2011, in both cases of GSR and NoGSR, the nodal RC values experience a decrease due to the HetBiv-correction. However, after normalisation, the NRC values, unanimously suggest an increase in nodal NRC values.
- Yeo17, in both cases of GSR and NoGSR, the nodal RC values suggest a decrease across all RSNs. After normalisation, in the case of GSR, there are only the Control, Default Modes and Saliency Ventral Attention that demonstrate significant increase in some of their nodes. The majority of other nodes across the remaining RSNs fail to show a significant increase. In case of NoGSR, all RSNs suggest an increase in their nodal normalised RCs.
- AAL2, in the case of GSR, all RSNs suggest a decrease in the nodal RC of their nodes. After normalisation, the nodal normalised RC values suggest a mixture of increase and decrease across nodes. Two RSNs, the Auditory and Sensorimotor, fail to demonstrate any significant changes in NRC values of any of their nodes. In the case of NoGSR, all nodes suggest a significant decrease in their nodal RC values. However, after normalisation, all nodes suggest a significant increase in their nodal normalised RC values.

5.3.5 Effect of Autocorrelation Correction on Graph Theoretical Measures of Density-Thresholding

Global Graph Theoretical Measures

We begin with the analysis of global measures (i.e. integration, segregation and complexity) over a range of densities. We examine the effect of correction methods by investigating the changes they may have caused in each of the aforementioned global measures at every density step. The changes are determined as a ratio of a specific graph theoretical value obtained from HetBiv-corrected networks over its corresponding value in a Naive-corrected network. Therefore, if the ratio is above one it suggests that the measure in HetBiv-corrected networks is higher compared to Naive-corrected networks. For each density step, the actual values of these two correction methods were tested for both cases of the GSR and Non-GSR aiming to ensure the changes are statistically significant.

Measures of integration (characteristic path length and its normalised version) are shown in first two rows of Figure 5.9. They suggest a decrease in level of integration of HetBiv-corrected networks across all four parcellation schemes and the choice of GSR. In the case of GSR, ICA100, Yeo17 and AAL2 the rates of change in characteristic path length for GSR networks are fairly similar (≈ 1.15), while, in Power2011, the rate of change is relatively lower (≈ 1.07). For the Non-GSR case, the growth in the rate of change is significant

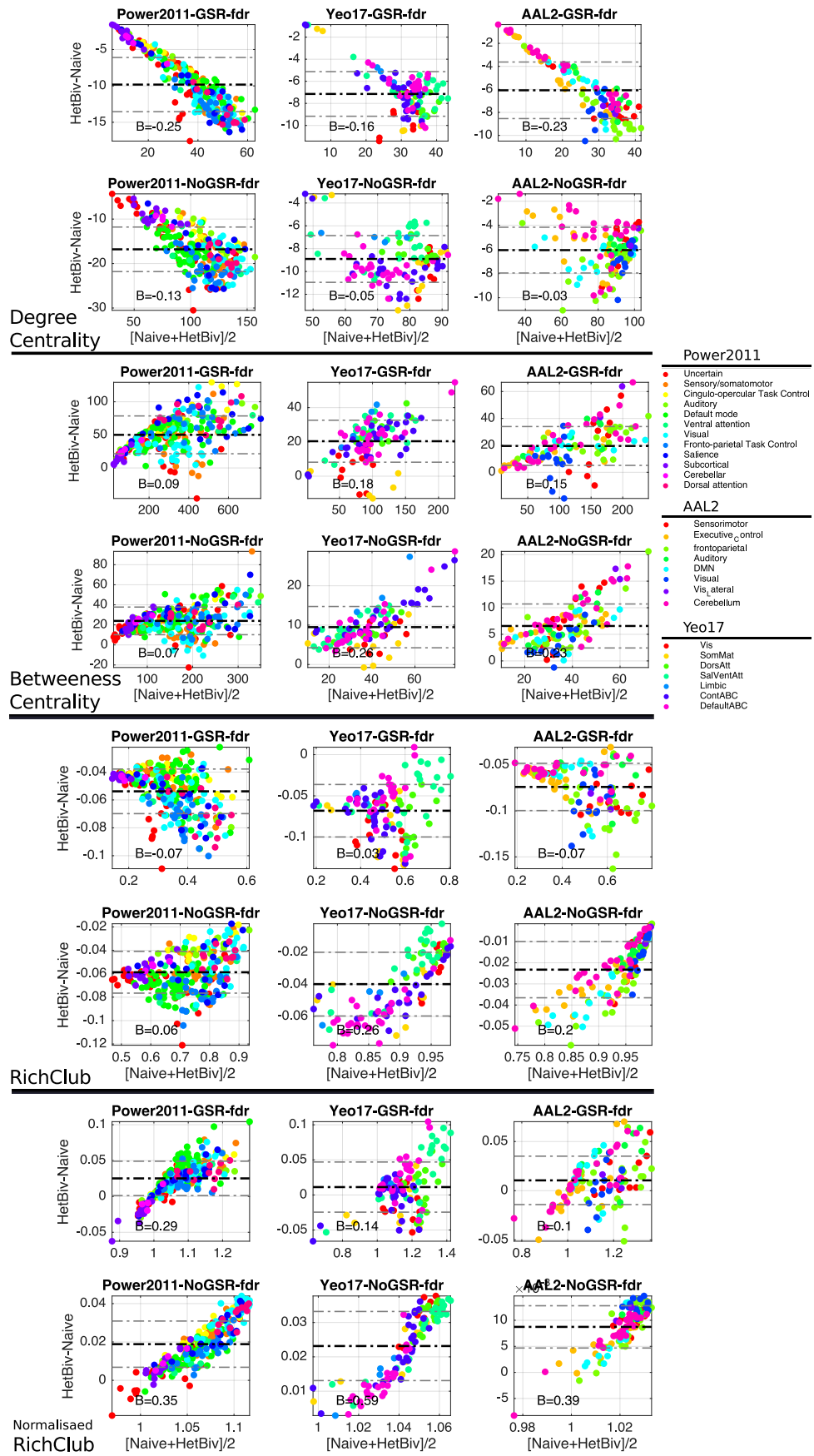


Figure 5.8: Disruption plots of the changes in the local measures (Degree Centrality, Betweenness Centrality and Rich Clubs) of the networks with statistical thresholding due to HetBiv correction. All of the x-axes represent the average of the measure across groups and all of the y-axes represents the differences in the values between the two groups.

across Yeo17, AAL2 and Power2011, however, the rate of change is no longer significant ICA100 beyond the cost-efficient densities. This implies that the effect of correction is more likely to be detected in sparser densities including the cost-efficient zone.

Measures of segregation (clustering coefficients, normalised clustering coefficient and modularity index) are shown on the third and fourth rows of Figure 5.9. In the case of GSR, contrary to characteristic path length, changes in the clustering coefficient is not consistent across densities. Broadly speaking, networks with GSR suggest a decrease in rate of change of clustering coefficients, although, this decrease is not always significant. On the other hand, in networks without GSR, the results suggest an increase in rate of change for Yeo17 and AAL2. Contrary to non-normalised clustering coefficient, the normalised version of clustering coefficient, Ω , demonstrates an increase in rate of change across the choice of GSR. The increase is consistent across densities. Similar to Path Length, as the density grows the rate of change approaches one which suggests that the difference between the HetBiv-correction and the Naive-correction becomes smaller. Another measure of segregation, the modularity index, also shows an increase in HetBiv-corrected networks for GSR networks. These changes are mainly significant in the sparse densities. In contrast, in the case of NoGSR, the modularity index fails to suggest a consistent change over densities.

Eventually, an increase in level of integration and segregation is reflected in the level of small-worldness. The small-worldness also shows an increase across the densities. Similar to the previous measure, in both cases of GSR and NoGSR, the changes are more significant in sparse densities and they become relatively smaller as the density grows until after a certain density, they are no longer significantly different.

Local Graph Theoretical Measures

In this section, we investigate the changes in local graph measures due to HetBiv-correction. For simplicity, we only investigate the changes on a common density between parcellation schemes and the choice of GSR. This common density is obtained by averaging across the cost-efficient density of each parcellation scheme with or without GSR. The common density is determined to be 15%. We discard ICA100 from further local analysis as the RSN assignment of each node is unknown. The disruption indices are summarised in Table 5.3 and Table 5.4 for the case of the GSR and NoGSR, respectively. Each disruption index corresponds to a RSN of a relative parcellation scheme which is later identified by different colours in Figure 5.10.

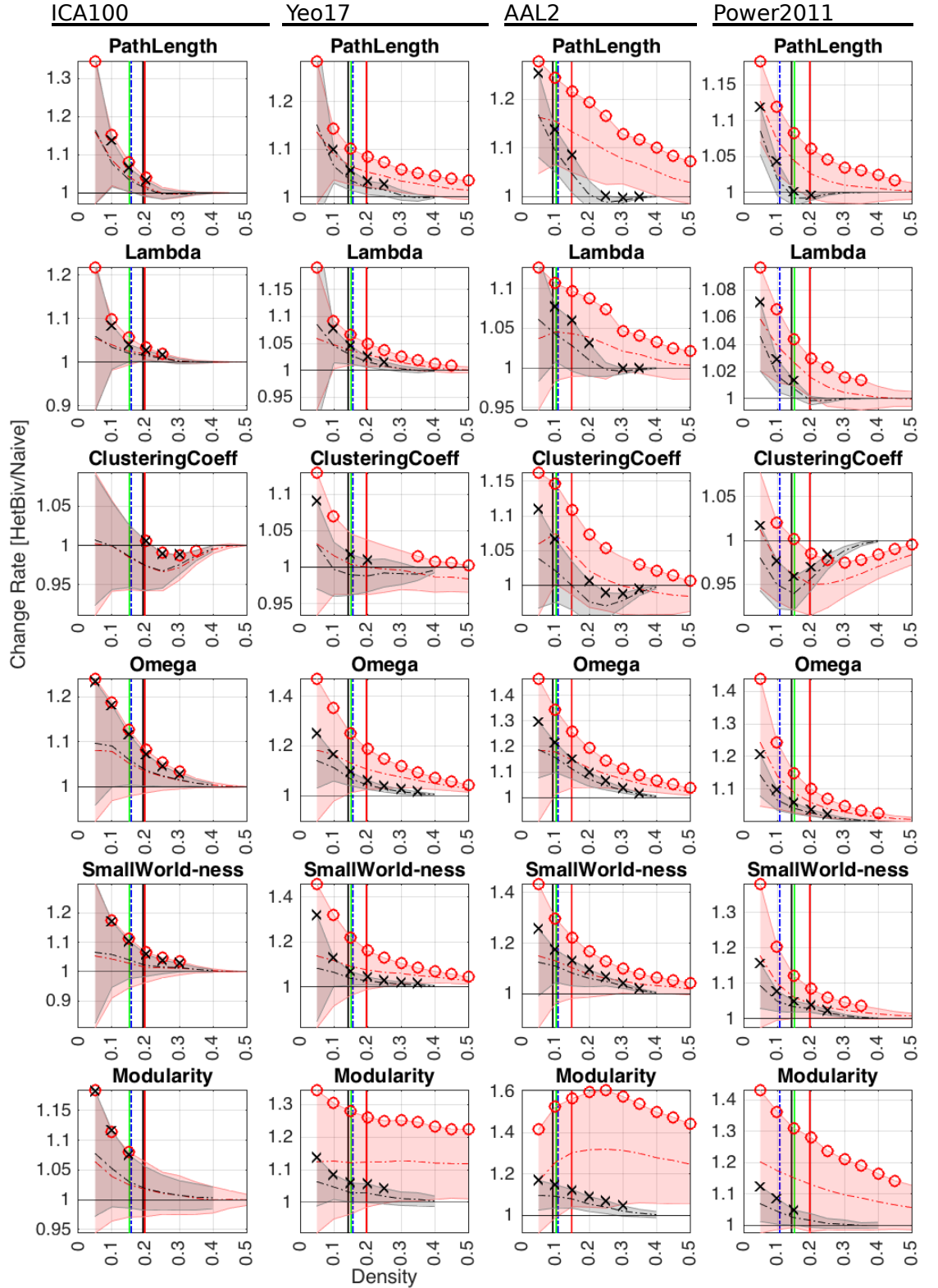


Figure 5.9: Change rates between global measure of network with Naive correction and HetBiv correction. Each column is a parcellation scheme and each row is global graph theoretical measure. Red circle indicates the statistical significance between the two correction in case of NoGSR. Black cross indicates the statistical significance between the two correction in case of GSR. Vertical lines show cost-efficient densities over different choices of GSR and correction methods. HetBiv-corrected, GSR (dashed blue); HetBiv-corrected, NoGSR (dashed green); Naive-corrected, GSR (solid black); Naive-corrected, NoGSR (solid red). All of the x-axis shows the density, and all the y-axis shows the Change Rate.

GSR - Density Thresholding

	Degree Centrality	Betweenness Centrality	Rich Club Coefficients	Normalised Rich Club Coefficients
Power2011				
Uncertain	-0.28	-0.29	-0.39	-0.27
Sensory/Somato	-0.31	-0.41	0	0.27
Cingulo Control	-0.11	0.05	0	0.41
Auditory	-0.16	-0.31	0	0.17
Default Mode	-0.22	-0.39	-0.11	0.32
Ventral Attention	-0.29	-0.37	0	0.22
Visual	-0.25	-0.45	-0.12	0.37
Fronto-parietal	-0.18	-0.37	-0.06	0.24
Saliency	-0.26	-0.34	-0.22	0.24
Subcortical	0.04	0.17	-0.26	-0.94
Cerebellar	-0.01	0.27	0	-0.45
Dorsal Attention	-0.01	-0.43	0	0.26
AAL2				
Sensorimotor	-0.26	0.16	-0.22	-0.2
Executive Control	-0.18	0.18	-0.29	0.62
Fronto-parietal	-0.32	-0.23	-0.19	-0.18
Auditory	-0.29	0.15	0.02	0.15
Default Mode	-0.4	0.22	-0.48	-0.24
Visual	-0.02	0.61	-0.1	0
Visual Lateral	-0.04	0.31	-0.04	0.04
Cerebellum	-0.07	0.14	-0.13	-0.44
Yeo17				
Visual	-0.15	-0.44	-0.12	0.27
Sensory	-0.18	-0.9	-0.31	-0.22
Dorsal Attention	-0.47	-0.33	-3.28	0.13
Ventral Attention	-0.13	-0.35	0.06	-0.06
Limbic	-0.06	0.9	-0.48	-0.97
Control	-0.22	-0.33	-0.19	-0.06
Default Mode	-0.27	0.02	-0.15	-0.23

Table 5.3: Disruption indices, as discussed in section 3.1.10, of RSNs of three atlases (Power2011, AAL2 and Yeo17) for the GSR case of local measures in Density Thresholding.

NoGSR- Density Thresholding

	Degree Centrality	Betweenness Centrality	Rich Club Coefficients	Normalised Rich Club Coefficients
Power2011				
Uncertain	-0.26	0.34	-0.32	-0.29
Sensory/Somato	-0.56	0.14	-0.51	-0.42
Cingulo Control	-0.05	0.2	-0.07	-0.42
Auditory	-0.24	0.23	-0.24	-0.54
Default Mode	-0.35	-0.07	-0.34	-0.37
Ventral Attention	-0.34	0.07	-0.26	-0.33
Visual	-0.39	-0.18	-0.2	-0.08
Fronto-parietal	-0.11	0.11	-0.05	-0.25
Saliency	-0.36	-0.02	-0.39	-0.39
Subcortical	0.56	0.92	0.2	-0.25
Cerebellar	0.06	0.90	-0.14	-0.58
Dorsal Attention	0.09	-0.08	0.07	-0.22
AAL2				
Sensorimotor	-0.37	0.59	-0.52	-0.91
Executive Control	-0.21	0.25	-0.26	-0.11
Fronto-parietal	-0.35	0.17	-0.34	-0.65
Auditory	-0.47	0.1	-0.38	-0.53
Default Mode	-0.27	0.08	-0.27	-0.36
Visual	-0.19	0.53	0.39	0.64
Visual Lateral	-0.09	0.03	-0.11	-0.27
Cerebellum	0.21	0.87	-0.09	-0.21
Yeo17				
Visual	-0.9	0.21	-0.06	0.17
Sensory	-0.25	0.43	-0.27	-0.17
Dorsal Attention	-0.34	-0.17	-0.35	-0.28
Ventral Attention	-0.07	-0.14	-0.04	-0.09
Limbic	0.8	0.95	0.42	0.12
Control	-0.35	-0.02	-0.35	-0.13
Default Mode	-0.35	0.16	-0.35	-0.59

Table 5.4: Disruption indices, as discussed in section 3.1.10, of RSNs of three atlases (Power2011, AAL2 and Yeo17) for the NoGSR case of local measures in Density Thresholding.

Degree Centrality

Changes in degree centrality show a negative change across all four parcellation schemes and GSR choices. The slope of the fitted line to these changes is close across choices of the GSRs. Also a remarkable proportion of the nodes experience a significant change due to the HetBiv correction. For networks in the case of GSR, these proportions are as 53% 78% 89% for Power2011, Yeo17 and AAL2, respectively. For NoGSR networks, these proportions are 53% 81% 85%, respectively.

- Power2011, in the case of GSR, suggests that the majority of RSNs demonstrate a significant mixture of an increase and decrease in the DC values. However, the Cingulo-opercular task control, Subcortical and Cerebral networks are the RSNs for which all of their nodes experienced an increase after HetBiv-correction. On the other

hand, all nodes of the Dorsal attention and Fronto-parital task control experience a decrease due to HetBiv-correction. It is worth noting that nodes of the Somatosensory system and the Fronto-parietal task control experience the highest negative association between their DC values and the changes due to HetBiv-correction. Salience and Fronto-parietal task control are among the RSNs, with several degree hubs, experienced the largest decrease. Degree centrality of networks without GSR, also, suggest that the nodes of the majority of RSNs experience both a decrease and increase in their DC values. Although, similar to GSR case, all significantly different nodes of Cingulo-opercular task control, Cerebral and Subcortical networks suggest an increase, whereas, the Fronto-parietal task control and the Dorsal Attention experience a decrease in their DC values.

- Yeo17, in the case of GSR, suggests that apart from Limbic system, all other RSNs experience a significant decrease in their DC values. Nodes of the Limbic system, unanimously, demonstrate an increase in their DC values due to HetBiv-correction. It is also worth noting that the Dorsal Attention network exhibits the largest positive association between the DC values and rates of change. In the case of NoGSR, similar to the case of GSR, the Limbic system is the sole network that suggests a consistent change across its nodes as they gained DC due to HetBiv-correction. Nodes of other RSNs shows a mixture of both an increase and decrease in their DC values, however, the association varies between them as the Dorsal Attention and Visual network shows the highest negative association, whereas, the Salience Ventral Attention network demonstrate the lowest association.
- AAL2, in the case of GSR, suggests that six out of eight RSNs exhibit a mixture of decrease and increase in their DC values, while, the two remaining networks, Visual and Visual lateral, show a consistent decrease in their DC values. It is also worth noting that more than 91% of the nodes in the Cerebellum show an increase. The Sensorimotor and Fronto-parietal network, which contain the degree hubs, experience the largest decrease. In the case of NoGSR, similar to the case of GSR, nodes of the Visual and Visual Lateral networks show an unanimous decrease in DC values of their nodes. Conversely, the nodes of the Cerebellum network show an increase in their DC values. The changes in other RSNs suggest a combination of increase and decrease. It is notable that the largest negative association can be seen in the Fronto-parietal and Auditory networks. Fronto-parietal network, which contains degree hub nodes, exhibits the largest decrease in its DC values.

Betweenness Centrality

Changes in Betweenness Centrality, are shown in the second row of Figure 5.10. In the case of GSR, 55%, 62% and 69% of the nodes in Power2011, Yeo17 and AAL2, respectively, are significantly affected by the HetBiv-correction. In the case of NoGSR, these proportions rise to 67%, 70% and 83%, respectively. The overall association between those changes and the BC values are not consistent across either the choice of parcellation scheme or the case of GSR. As for Power2011-GSR and Yeo17-GSR this association is positive, while, for remaining conditions the results suggest a close-to-zero or negative association.

- Power2011, in the case of GSR, suggests that all nodes of the Cingulo-opercular task control, Cerebellum and Subcortical networks experience a significant increase in their BC values. The rest of the RSNs, experience both an increase and decrease across their nodes. It is also worth mentioning that the Auditory, Default Mode and Fronto-parietal networks are among the RSNs for which the BC hubs experienced the largest decrease due to HetBiv-Correction. In the case of NoGSR, the Subcortical, Cerebral and Cingulo-opercular task control, Dorsal Attention, Fronto-parietal task control, Auditory and Ventral Attention networks are among the networks which unanimously experience an increase. The remaining networks such as the Default Mode and Salience networks fail to suggest a consistent change, neither an increase nor decrease, in their BC values.
- Yeo17, in the case of GSR, suggests that nodes of the majority of RSNs experience both an increase and decrease. However, the Limbic system exhibits a consistent increase in its node BC values. It also suggest the largest positive association, while, the largest negative association is suggested by the Visual networks. Salience Ventral Attention network, which contain the BC hubs, experienced the largest decrease. In the case of NoGSR, the majority of networks such as the Limbic, Visual, Somatosensory, Visual and Default Mode networks suggest a unanimous increase in their BC values, while, the remaining networks suggest a mixture of increase and decrease across their nodes. Similar to the case of GSR, the Limbic system suggests the highest positive association. The majority of the BC hubs which belong to the Default Mode network suggest an increase in their BC values due to HetBiv-correction.
- AAL2, in the case of GSR, suggests that all of the RSNs demonstrate a consistent increase in their BC values, however, the small portion of nodes within the Fronto-parietal network suggest a mixture of increase and decrease in their BC values. The Visual network has the largest positive association among the RSNs. It is also worth noting that the effect of HetBiv-correction is not consistent across the RNS as BC

hubs in the Cerebellum experience a relatively large increase in their BC value, while the BC hubs of the Fronto-parietal networks suggest the opposite. In the case of the NoGSR, all RSNs show a consistent increase in their BC values across the nodes. The Cerebellum demonstrates a remarkable positive association.

Normalised and non-normalised nodal Rich Club coefficients

Normalised and non-normalised versions of the Nodal Rich Club coefficients are the last graph measure that we use to investigate the changes due to HetBiv-correction. These changes are summarised in last two rows of Figure 5.10. In the Nodal Rich Club measures, in the case of GSR, the percentage of nodes that were found to be significantly affected are 15%, 50% and 70% for Power2011, Yeo17 and AAL2, respectively. However, these percentages rise to 54%, 78% and 90% in the case of NoGSR. These percentages even become larger in the case of normalised Rich Club for both GSR (98% 91% 65%) and NoGSR cases (95% 89% 94%).

- Power2011, in the case of GSR, suggests only a small number of nodes (mostly from the Default Mode, Visual and Fronto-parietal networks) exhibit a significant decrease in their nodal RC coefficients. Conversely, the result of changes in the normalised RC Coefficients suggest that all of the nodes across the 12 RSNs show a level of significant increase in their nodal normalised RC coefficients. The Dorsal attention, Default Mode and Subcortical networks are among the RSNs which show the highest level of increase. However, it is notable that the majority of the nodes of the Subcortical network were not contributing to the RC organisation since their RC coefficients were not identified to be significantly different. In the case of NoGSR, changes in the nodal RC coefficients of nodes across RSNs are not consistent although, the number of nodes which experienced significant changes in their nodal RC coefficients is remarkably larger. The Subcortical, Cerebral, Auditory and Cingulo-opercular task control networks demonstrate a unanimous increase in their nodal RC coefficients, whereas, the changes across remaining networks are a combination of increase and decrease. After normalisation, all RSNs demonstrate an increase in their nodal RC coefficients. However, all of the nodes in the Cerebral and Subcortical networks fail to be considered as a part of a RC organisation. The nodes which significantly experienced an increase due to HetBiv-correction are mostly in the Fronto-parietal task control, Dorsal Attention and Default Mode networks.
- Yeo17, in the case of GSR, suggests that the Limbic system, with all its nodes increased, and the Visual network, with all its nodes decreased, are two extremes of

HetBiv-correction effects. The remaining RSNs experience a combination of increase and decrease in their nodal RC coefficients. On the other hand, the nodal normalised RC coefficients of nodes across all RSNs consistently experience an increase. The Salience and Default Mode networks are among the RSNs with the highest nodal normalised RC coefficient values which also suggest an increase due to HetBiv-correction. Despite the Limbic system suggesting a relatively large association between its normalised RC coefficient, it fails to contribute to the RC organisation. In the case of NoGSR, Limbic system is the only network which shows a unanimous increase in nodal RC coefficients, while, the effect of HetBiv-correction on nodal RC coefficients of nodes across other RSNs are shown to be mixture of increase and decrease. Contrary to the results of the RC coefficients, the normalised RC coefficients across all RSNs, consistently, suggest an increase in their normalised RC coefficients. It is worth noting that the largest negative association between nodal RC coefficients and the changes due to HetBiv-correction is suggested by the Default Mode network.

- AAL2, in the case of GSR, suggests that the RC coefficients of nodes within the Visual, Visual Lateral and Auditory networks experience a consistent decrease, while, nodes of the remaining RSNs experience a combined increase and decrease in their RC values. However, after normalisation all RSNs, unanimously, demonstrate an increase in their nodal normalised RC coefficients. Contrary to rest of the RSNs, the Executive Control and Cerebellum networks suggest a strong negative association between their nodal normalised RC coefficients and the changes due to HetBiv-correction. In the case of NoGSR, the Visual and Visual lateral networks still suggest a unanimous decrease, however, the remaining networks demonstrate an inconsistency in changes to their RC coefficients. However, after the normalisation, similar to the GSR case, all the RSNs suggest a significant increase for nodes across all RSNs. Despite a unanimous increase across the RSNs the Auditory network suggests a strong negative association between its nodal normalised RC coefficients and changes due to HetBiv-correction. This association in the Visual network suggests the opposite. Finally, as one may notice, the results in Figure ??, the Power2011 shows more systematic changes compared to other atlases. This association is fuelled by, firstly, the same-size ROIs in Power2011 and, secondly, the heterogeneous effect of the autocorrelations.

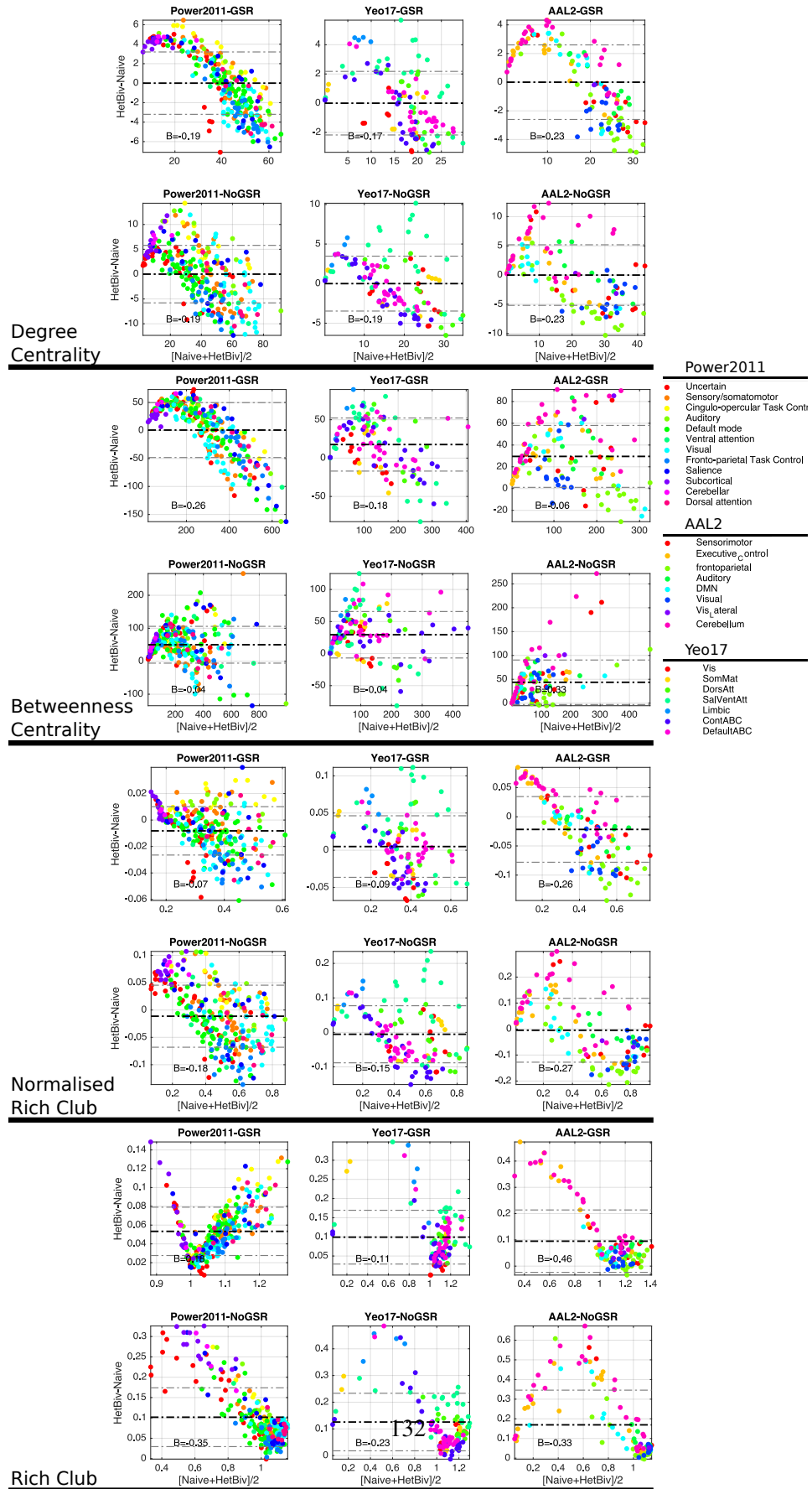


Figure 5.10: Disruption plots of changes in local measures (Degree Centrality, Betweenness Centrality and Rich Clubs) of network with density thresholding due to HetBiv correction. All of the x-axes represent the average of the measure across groups and all of the y-axes represents the differences in the values between the two groups. 132

5.4 Discussion

Autocorrelated BOLD signals, are one of the many issues that accompanies the analysis of the functional connectome of the human brain. The autocorrelated BOLD signals can, often severely, inflate the degree of freedom due to independency between observations. This violates the presumption of the majority of statistical testing methods which are centred on the data-points being completely independent. In this chapter, we discuss the effect of autocorrelated BOLD signals on topological properties of the functional connectome. In more detail, we show that autocorrelation does not affect all of the three main parts of the brain, White Matter, Grey Matter and CSF, equally. Grey matter is shown to be more affected by the non-white noises. Furthermore, we also showed that the distribution of correlation lengths suggests that the effect of the autocorrelation is not uniform across grey matter itself, but rather, it suggests a level of heterogeneity across different parts of the brain. Therefore, it is crucial to investigate the effect of autocorrelation on parcellation schemes which are commonly used to decimate the raw functional connectome to a smaller number of either anatomically or functionally meaningful regions. Our findings suggest that there is a strong association between the volume of ROIs and correlation length since the parcellation schemes with larger ROIs are more vulnerable to the effect of autocorrelation.

The existing methods for mitigating the effect of autocorrelation in functional connectivity analysis consider the BOLD signals as a first-order autoregressive process. However, it is widely known that the BOLD signals are constructed of far more complicated temporal processes. In this work, we propose a novel method of correcting the degree of freedom for Pearson's correlation coefficients which is based on estimating the variance of the correlation coefficients. The proposed methods examine the temporal structure of BOLD signals beyond the first-order autoregressive process by employing Bartlett's 'general process'. This technique considers the autocorrelation function of the all temporal lags available for a time series. In addition to considering several temporal lags of an ACF, the method considers a more complicated autocorrelation structure constructed by combining the ACF of both time series involved in the correlation. The bivariate nature of the method helps us to account for the spatial heterogeneity of the autocorrelation effect present in the cortical and sub-cortical regions of the brain. The bias due to the inner product of the ACFs is also corrected by extra steps which was discussed in section 5.1.7. However, by validating the methods on real-data, the results suggest that the biased version still has fairly close results to the unbiased version, which motivates the case of computational efficiency. In traditional applications of the BCF general-process, it is suggested that accounting for the ACF of all temporal lags may result in meaningless estimation of very distant lags (or larger lags, see Eq. 5.9) where the autocorrelation is likely to be zero. To mitigate this

biased estimation, it is commonly advised to curb/shrink the autocorrelation function so that only a few early lags remain in the estimation. The proposed method demonstrates that it is unnecessary to use any form of arbitrary curbing or shrinking strategies as, for long time series, the probability of the occurrence of negative autocorrelation is as large as positive autocorrelation, and therefore, apart from early autocorrelation lags, the remaining lags cancel out each other.

To compare the HetBiv method with other existing methods, we use a real-data technique to avoid oversimplifying the BOLD signals. In traditional simulation studies, it is often common to generate time series with only one or very few orders of the autoregressive function. Instead, we exploit the realisation that the BOLD signals are irreproducible, and therefore, the correlation coefficient between one randomly-picked time series of a subject and another randomly-picked time series of another subject is expected to be zero. After iterating the inter-subject correlation of randomly-picked nodes for a large number of realisations, the obtained distribution of correlation coefficients is expected to be a univariate mean-zero normal distribution. Any distribution with a different mean or variance suggests a systematic confound due to autocorrelation that exists in the BOLD signals. By comparing the shape and tail of each distribution obtained from each method against the ground truth, we found that the methods of the HetBiv family outperform the other methods discussed in this chapter. It is worth mentioning, that the FSLnet implementation for estimating the variance of the correlation coefficients is expected to out-perform the other AR(1)-based methods if the number of simulations goes beyond the number of nodes. However, it is worth noting that the running of thousands of synthetic time series for each pair of the nodes still seems to be computationally cumbersome.

Using the HetBiv method to estimate the effective degree of freedom, can not only drastically affect the hypothesis testing required for statistical thresholding, but it can also affect the density thresholding in the sense that the previous correlation coefficients are heterogeneously changed during the standardisation process. In this study, we investigate the effect of HetBiv-correction on both thresholding methods with and without global signal regression.

In Statistical Thresholding, global signal regression reduces the variability between global densities of the functional connectome of subjects, while the HetBiv-correction reduces their values. Bearing in mind that larger variation between the density of networks may introduce a bias in analysis of graph theoretical measures due to differences in the average degree of nodes, global signal regression is shown to be a more reliable option for Statistical Thresholding in graph analysis since it minimises the inter-subject variation in global densities. A reduction in the density of each subject implies that there will be a decrease in the Degree Centrality of the nodes, however, it was shown that this reduction

does not homogeneously affect all nodes. Analysis of Degree Centrality suggests a decrease in almost all of the resting-state functional networks, however, the magnitude of this reduction is not similar across the RSNs. Broadly speaking, the Default Mode network, and the RSNs related to the Fronto-parietal area of the brain (Fronto-parital task control in Power2011, Fronto-parietal in AAL2 and SalVenAtt and parts of the Control network in Yeo17) are the networks which are shown to be largely affected by the HetBiv-correction, whereas, the Subcortical regions (Subcortical in Power2011, Limbic in Yeo17) along with Cerebellar networks (Cerebral in Power2011, Cerebellum in AAL2) are among the least affected RSNs. Despite the reduction in the Degree Centrality, all of the RSNs experience an increase in their Betweenness Centrality, consistently, across all parcellation schemes. Surprisingly, nodes from the Default Mode and Fronto-parietal networks are among the BC hubs which experienced the largest increase in their BC values, in spite of the reduction in their the DC values. Analysis of changes in DC and BC values of the Cerebellum and Subcortical networks also suggest a small, but still significant, increase in the values. The changes in Rich Club coefficients before and after normalisation are not consistent as in non-normalised version of the RC coefficients, there is a unanimous decrease across all RSNs, however, after normalisation, the results suggest an increase in the nodal normalised RC coefficients. The initial reduction of the non-normalised RC coefficients is likely to be the result of the decrease in the average degree of nodes due to the HetBiv-correction in statistical thresholding, although the normalised versions suggest that the initial increase may help reveal the more tightly connected RC organisations among the degree hubs. Further to the effect of the HetBiv-correction on the local measures, we also investigate the effect of this correcting of the degree of freedom on global measures. Analysis of the Clustering coefficient suggests a decrease due to HetBiv correction. However, after normalisation, normalised clustering coefficients, Ω , indicate an increase in the level of segregation. This increase is also notable in the modularity indexes. The characteristic path length and its normalised version, Λ , indicates an increase after the HetBiv-correction. However, normalised characteristic path length is fairly close to one, indicating the level of integration in the networks are still as high as a random graph. Changes in the level of segregation and integration results in an increase in the level of small-worldness in the networks.

Contrary to Statistical Thresholding, in the Density Thresholding, the density of networks is fixed between the HetBiv and Naive corrections. Therefore, tracking the heterogeneity in changes, specifically in the Degree Centrality, is more straightforward as the bias due to the difference in network densities does not longer exist. The Subcortical and Cerebral regions are among the networks which experience an increase in their DC values, however, this increase is compensated by a reduction in DC values of the Visual, Fronto-parietal and Dorsal Attention networks. Other networks such as the Default Mode and

Auditory networks suggest a balance between increase and decrease. The Betweenness Centrality of nodes within the Cerebral and Subcortical networks suggest a slight increase. The changes in Betweenness Centrality of other networks are not consistent across either parcellation schemes or choice of GSR. For example, The BC of all networks shows an increase in AAL2 while, in other parcellation schemes the changes in BC values is a mixture of increase and decrease. Similarly, the Rich Club coefficients of the Subcortical and Cerebral network show an increase while the RC coefficients across the remaining RSNs fail to demonstrate consistent changes across the parcellation schemes. Although the nodal RC coefficients fail to exhibit a consistent change across the RSNs of different parcellation schemes, the normalised nodal RC coefficients of all parcellation schemes suggest a consistent increase. We also investigated the effect of HetBiv-correction on the global measures. Changes in the Clustering coefficient over the densities fails to demonstrate a consistent significant difference in the level of segregation. However, almost all parcellation schemes suggest a significant increase around the cost-efficient densities. It is also worth noting that over the majority of density steps the Clustering Coefficient shows a decrease due to HetBiv-correction. However, the difference approaches zero as the density grows until the Clustering coefficients in the Naive and HetBiv-correction become identical. Normalised Clustering coefficients, Ω , show a consistent increase across the densities due to HetBiv-correction for both GSR and Non-GSR choices. Similarly, Modularity index also suggests a significant increase due to HetBiv-correction for GSR and NoGSR cases. Eventually, these changes suggest an increase in the level of segregation for networks for which their degree of freedom were corrected by HetBiv method. The Characteristic Path Length and its normalised version, Λ , also suggest an increase due to HetBiv correction, however, the magnitude of the difference is greater for NoGSR cases. A statistically significant increase in Λ suggests a decrease in the level of integration. Eventually, these increases in the level of integration and segregation are reflected later in the measure of small-worldness and suggest an increase in level of complexity of the networks due to HetBiv-correction. It is important to note that despite the increases in the normalised characteristic path length, the values are still fairly close to one, which suggest that the HetBiv corrected networks still maintain an integration level as large as a random network.

The results from Density and Statistical thresholding confirm that the HetBiv correction affects the Fronto-parietal regions, Subcortical, Cerebral, Visual networks in various ways. The increase in the Subcortical, Cerebral and Visual networks may be relevant to relatively smaller effects of autocorrelation in these regions, which leads the HetBiv-correction to have a smaller effect, and consequently, leads to the detection of more edges in the aforementioned areas. Also the decrease in the Fronto-parietal, Default Mode and Dorsal Attention networks suggest that it is more likely that the inflated degree of freedom in these

areas due to large effect of autocorrelation may lead the edges in these areas to falsely show a strength in Naive-corrected networks. Regardless of the choice of parcellation scheme or GSR, the results suggest that HetBiv-correction reveals a higher level of segregation and a lower rate of integration. As a rate of growth in segregation is higher than the integration, and integration level is still fairly similar to that for random networks, the complexity level of the networks is suggested to be increased due to HetBiv-correction.

Among the parcellation schemes, the Power2011, exhibits relatively robust results against the autocorrelation since the analysis of the shape and tails of the null distribution, obtained from the comparison techniques, suggest a higher similarity between its inter-subject correlation coefficients and the ground-truth. Further, analysis of the results from the Density Thresholding suggest that the Power2011 was relatively less affected by the HetBiv correction in both the segregation and integration measures. In terms of GSR, Power2011, Yeo17 and AAL2 show changes in their graph theoretical measures after and before the GSR, whereas, the results of ICA100 do not demonstrate a significant change due to GSR.

5.5 Summary

In this chapter we focused on the impact of non-white noises in the analysis of functional connectome of the human brain. We also proposed a series of methods which can help mitigating this effect.

- Through ROI-based analysis of correlation length, we demonstrated that the autocorrelation (non-white noises) are rather distributed heterogeneously across the brain surface. This findings contradicts with the previous assumptions of the dependencies across the BOLD signal time-points.
- By investigating the correlation lengths of wide range of parcellation schemes, we also showed that the impact of autocorrelation is strongly related to the size of the ROIs in a parcellation scheme. As schemes with fixed size of ROIs (such as CC200 and Power) shows lower level of autocorrelation compared to other schemes (such as AAL, AAL2 and Yeo17).
- To mitigate this effect, we proposed a three bivariate heterogeneous methods, called the Het-Biv. These methods, contrary to other existing methods, accounts for spatial heterogeneity as well as complicated temporal structure of the BOLD signals.
- We used inter-subject scrambling method, assuming the in the ideal case the correlation between two BOLD signals of two different subjects should follows univariate normal distribution. We used Kolmogorov-Smirnov to quantify the performance of the

proposed methods. The results suggest the HetBiv methods demonstrate the smallest KS statistics (≈ 0.003 for the HCP cohort), and therefore, is more successful in accounting for the effect of autocorrelation. As the KS statistics are not sensitive to the tails of a distribution, we separately investigated the tails using false-positive rates. The results suggest that the HetBiv methods outperforms the existing methods in controlling the tail of the distribution of correlation coefficients which are crucial in further statistical inferences.

- Finally, we investigated the effect of this correction for both cases of statistical and density thresholding. The results suggest significant changes in topological description of the functional connectome.

Chapter 6

Conclusions and Future Work

Aiming to deliver the four objectives discussed in section 1.1, we developed a series of approaches and methods for multi-subject inference of graph theoretical methods. We also investigated the effects of noise on the topological description of functional connectivity by proposing a novel method aiming to mitigate the effect of non-white noise.

In chapter 3, we presented a general overview about how existing graph theoretical methods can be used to confirm the previous findings about the topological features of the functional connectivity of the human brain. We proceeded this chapter by showing how these methods can be used as bio-markers to detect the changes in the functional connectome in patients with schizophrenia. Aligned with the objectives, a statistical approach was proposed (see section 3.1.9) to facilitate between-group statistical inference on RC coefficients and RC organisation between two groups of healthy control and patient. Previously, RC analysis were performed on group-averaged networks which result in loss of the within-group variabilities, however the proposed method integrates the changes in RC coefficients and RC organisations across subjects and therefore makes hypothesis testing easier.

In chapter 4, we elaborated the notion of a Rich Club by proposing a novel method for detecting highly connected nodes of a network, named the Rich Block. This method can be seen as a complimentary approach to SBM in the sense that it utilises the SBM estimated parameters to facilitate a group summarisation of highly connected network cores. Specifically, by using the Het-SBM, the RB can be helpful in between-group inference of these cores since the nodes' block assignments remain identical for each subject. Further to group inference, the RB also accounts for the profile of connectivity as well as subject-level covariates (see section 4.1.3). For detection of RB organisations, we developed an approach to produce randomised networks, called a Tensor Null Model (see section 4.1.3). The TNM not only enables detection of a RB organisation, it is also remarkably faster (63k times) than traditional randomisation techniques as each of the one thousand networks takes

$\approx 0.5s$, while, the TNM only requires a execution time as long as a vector multiplication. Finally, we showed that how the series of RB methods can be used to detect changes in the topological properties of patients with schizophrenia.

In chapter 5, we discussed how the topological properties, discussed in the two aforementioned chapters, can be affected by correlated noise in BOLD timeseries. We investigated the overall effect of this noise across different parcellation schemes and concluded that the effect is not spatially homogeneous. We proposed a method, named HetBiv, to estimate the inflation in sampling variance due to this class of noise. Additionally, contrary to traditional methods which consider the effect to be spatially homogeneous across the brain, HetBiv estimates the sampling variance for every connection which accounts for spatial heterogeneity in the effect of correlated noise. By validating the HetBiv-correction by a scrambled inter-subject BOLD time series (see section 5.1.9), we concluded that the method outperforms existing methods by more accurate estimation of sampling variance. In addition, we compared the effect of the correction of graph theoretical measures and demonstrated that ignoring the effect of correlated noise may introduce a bias in the topological description of the functional connectivity.

The description of brain connectivity presented in this work was merely centred on describing the functional activities. However, recent findings suggest that brain function is rather derived from the structure of the human brain (Sporns, 2013). Therefore, combining the different imaging modalities in the analysis of the brain connectomics appears to be essential in gaining a wider insight into the function-structure coupling of the brain. Several new brain imaging repositories such as the Human Connectome Project (HCP) bring high quality functional and structural images (Van Essen *et al.*, 2012). In addition, newly proposed statistical models such as Het-SBM, discussed in section 4.1.2, Multilayer stochastic and deterministic clustering methods can facilitate integration between functional and structural connectivity (Kivelä *et al.*, 2014; Valles-Catala *et al.*, 2016; Mucha *et al.*, 2010).

The work presented in this thesis relied upon the presumption that brain activities are stationary during the period of a scan. However, it is crucial to highlight that the recent investigations of fMRI and EEG data show that the brain states can be altered during the period of a scan (Hutchison *et al.*, 2013). Therefore, studying the brain functional connectivity as a dynamic network appears to be essential. Dynamic connectivity of the human brain was also shown promising evidence that graph theoretical measures, such as modularity, can be helpful bio-makers in order to gain a better understanding of the learning process in the human brain (Bassett *et al.*, 2011). Summarising the majority of graph theoretical measures in a dynamic setting were proven to be cumbersome. However, by considering former ‘subjects’ as new ‘time-points’ in the analysis of Het-SBM and a Rich Block we can estimate a common block assignment among the time-points, and therefore,

examine the changes in network cores over a period of time using Rich Block coefficients and organisation.

Finally, motivated by emerging large-scale, publicly available, datasets which offer high quality human phenotypes ([Collins, 2012](#)), the graph theoretical measures presented in this work can be used as tools to form a linkage between behaviours, demographics and life-styles and topological properties of the human brain ([Mišić and Sporns, 2016](#)). In addition, these phenotypes can also be used in investigations of disease progression and the effects of exposures in populations ([Smith *et al.*, 2015](#); [Dubois and Adolphs, 2016](#)).

Appendix A

Anatomy of the Human Brain

Anatomy and Functional Areas of the Brain

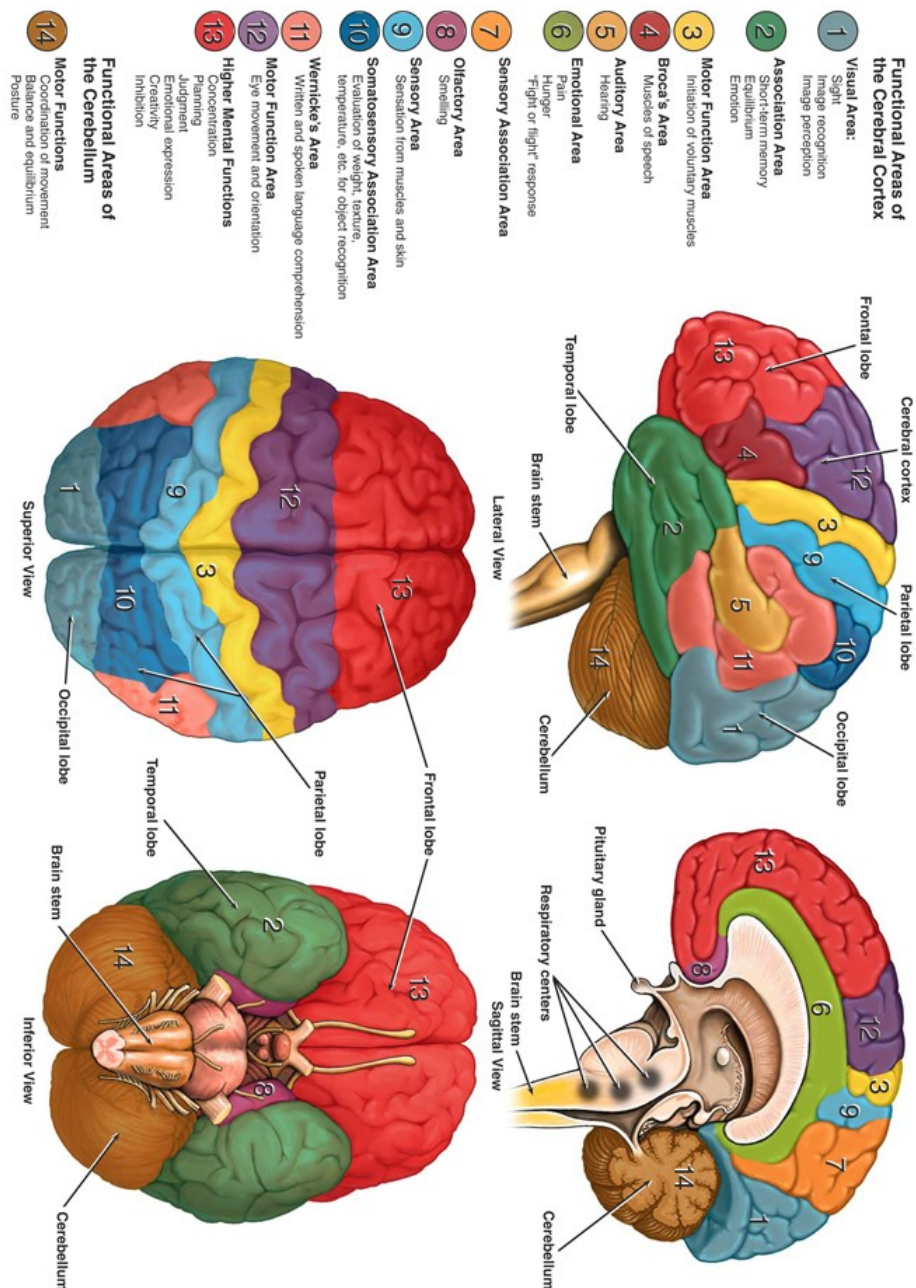


Figure A.1: illustrates the anatomy of the human brain including the lobes and the areas within each lobe.

Appendix B

Rich Block

B.1 Derivation of the Variational Bound

In this section, we derive a lower bound for the incomplete data, represented by the marginal density $f(\mathbf{x}; \boldsymbol{\pi}, \boldsymbol{\alpha})$. Setting a sample space of \mathbf{Z} as \mathcal{S} , we can derive the following inequality

$$\begin{aligned}\log f(\mathbf{x}; \boldsymbol{\pi}, \boldsymbol{\alpha}) &= \log \left[\sum_{\mathbf{z} \in \mathcal{S}} f(\mathbf{x}, \mathbf{z}; \boldsymbol{\pi}, \boldsymbol{\alpha}) \right] \\ &= \log \left[\sum_{\mathbf{z} \in \mathcal{S}} f(\mathbf{x}, \mathbf{z}; \boldsymbol{\pi}, \boldsymbol{\alpha}) \frac{f^*(\mathbf{z}; \boldsymbol{\tau})}{f^*(\mathbf{z}; \boldsymbol{\tau})} \right] \\ &= \log \left[\mathbb{E}_{f^*} \left(\frac{f(\mathbf{x}, \mathbf{Z}; \boldsymbol{\pi}, \boldsymbol{\alpha})}{f^*(\mathbf{Z}; \boldsymbol{\tau})} \right) \right] \\ &\geq \mathbb{E}_{f^*} \left(\log \left[\frac{f(\mathbf{x}, \mathbf{Z}; \boldsymbol{\pi}, \boldsymbol{\alpha})}{f^*(\mathbf{Z}; \boldsymbol{\tau})} \right] \right) \quad (\text{by Jensen's inequality}) \\ &= \mathbb{E}_{f^*} \left(\log[f(\mathbf{x}, \mathbf{Z}; \boldsymbol{\pi}, \boldsymbol{\alpha})] \right) - \mathbb{E}_{f^*} \left(\log[f^*(\mathbf{Z}; \boldsymbol{\tau})] \right).\end{aligned}\tag{B.1}$$

In particular, the lower bound in Eq. (B.1) is closely related to the Kullback-Leibler (KL) divergence of $f^*(\mathbf{z}; \boldsymbol{\tau})$ to $f(\mathbf{z}|\mathbf{x}; \boldsymbol{\pi}, \boldsymbol{\alpha})$, as

$$\begin{aligned}\text{KL} \left[f^*(\mathbf{z}; \boldsymbol{\tau}) \parallel f(\mathbf{z}|\mathbf{x}; \boldsymbol{\pi}, \boldsymbol{\alpha}) \right] &= \log f(\mathbf{x}; \boldsymbol{\pi}, \boldsymbol{\alpha}) \\ &\quad - \left[\mathbb{E}_{f^*} \left(\log[f(\mathbf{x}, \mathbf{Z}; \boldsymbol{\pi}, \boldsymbol{\alpha})] \right) \right. \\ &\quad \left. - \mathbb{E}_{f^*} \left(\log[f^*(\mathbf{Z}; \boldsymbol{\tau})] \right) \right].\end{aligned}\tag{B.2}$$

Rearranging the last equation, we write the lower bound Eq. (B.1) instead as

$$\log f(\mathbf{x}; \boldsymbol{\pi}, \boldsymbol{\alpha}) - \text{KL} \left[f^*(\mathbf{z}; \boldsymbol{\tau}) \parallel f(\mathbf{z}|\mathbf{x}; \boldsymbol{\pi}, \boldsymbol{\alpha}) \right].\tag{B.3}$$

This shows that the lower bound in Eq. (B.1) is precisely attained when the KL divergence is zero, that is, when $f^*(z; \tau)$ coincides with $f(z|x; \pi, \alpha)$. As we cannot explicitly calculate $\log f(x; \pi, \alpha)$ due to the intractability of $f(z|x; \pi, \alpha)$, we find a family of densities with variational parameter τ to approximate $\log f(x; \pi, \alpha)$ by its lower bound,

$$\begin{aligned} \log f(x; \pi, \alpha) - \text{KL}\left[f^*(z; \tau) \parallel f(z|x; \pi, \alpha)\right] \geq \\ \mathbb{E}_{f^*}\left(\log[f(x, Z; \pi, \alpha)]\right) \\ - \mathbb{E}_{f^*}\left(\log[f^*(Z; \tau)]\right). \end{aligned} \quad (\text{B.4})$$

To indicate the role of τ , the lower bound is formally denoted as

$$\begin{aligned} \mathcal{J}(f^*(z; \tau); \pi, \alpha) = \mathbb{E}_{f^*}\left(\log[f(x, Z; \pi, \alpha)]\right) \\ - \mathbb{E}_{f^*}\left(\log[f^*(Z; \tau)]\right). \end{aligned} \quad (\text{B.5})$$

In particular, the natural option for $f^*(z; \tau)$ is a categorical distribution with block-specific probabilities, independent for each node:

$$f^*(z; \tau) = \prod_{i=1}^n \prod_{q=1}^Q \tau_{iq}^{z_{iq}}, \quad (\text{B.6})$$

where $\sum_{q=1}^Q \tau_{iq} = 1$. Note, that this gives the probability of node V_i belonging to block q and as $\mathbb{E}_{f^*}(Z_{iq}) = \tau_{iq}$ and that $\mathbb{E}_{f^*}(Z_{iq}Z_{jl}) = \tau_{iq}\tau_{jl}$ the expected terms in Eq. B.5 are simple to compute.

B.2 Stochastic Blockmodel

Given the block assignment Z and given that the vertex V_i is a member of block q , the conditional distribution of the degree of node V_i can be written as

$$\begin{aligned} \rho(V_i|Z_{iq} = 1) \sim \text{Binomial}(n-1, \bar{\pi}_q) \approx \text{Poisson}(\lambda_q), \\ \bar{\pi}_q = \sum_{\ell=1}^Q \hat{\alpha}_\ell \hat{\pi}_{q\ell} \quad \lambda_q = (n-1)\bar{\pi}_q, \end{aligned}$$

where $\bar{\pi}_q$ is a weighted sum of connection rates between block q and the rest of the network and λ_q is the expected degree of block q . To show that this, we first need to consider

probability of updating a degree of a node $\rho(V_i)$, that is

$$\begin{aligned}
P(X_{ij} = 1|Z_{iq} = 1) &= \frac{P(X_{ij} = 1, Z_{ij} = 1)}{P(Z_{iq} = 1)} \\
&= \frac{\sum_{l=1}^Q P(X_{ij} = 1, Z_{iq} = 1, Z_{jl} = 1)}{P(Z_{iq} = 1)} \\
&= \frac{\sum_{l=1}^Q P(X_{ij} = 1|Z_{iq} = 1, Z_{jl} = 1)P(Z_{iq} = 1, Z_{jl} = 1)}{P(Z_{iq} = 1)} \\
&= \frac{\sum_{l=1}^Q P(X_{ij} = 1|Z_{iq} = 1, Z_{jl} = 1)P(Z_{iq} = 1)P(Z_{jl} = 1)}{P(Z_{iq} = 1)} \\
&= \sum_{l=1}^Q P(X_{ij} = 1|Z_{iq} = 1, Z_{jl} = 1)P(Z_{jl} = 1) \\
&= \sum_{q=1}^Q \hat{\alpha}_l \hat{\pi}_{ql} = \bar{\pi}_q.
\end{aligned}$$

Finally, as the edges in the SBM are assumed to be conditionally independent given their cluster labels, we can see that the $n - 1$ independent Bernoulli random variables (i.e. the maximal degree of any node is $n - 1$ as a node cannot be contacted to itself) can be approximated by the Binomial density with parameter $\bar{\pi} = (\bar{\pi}_1 \dots \bar{\pi}_Q)$. By applying Poisson approximation to the Binomial distribution and a suitable adjustment of its parameter $(n - 1)\bar{\pi} = \lambda$.

The expected connection rate between classes q and l is given as

$$\mathbb{E}(E_{ql}) = \frac{n(n - 1)\hat{\alpha}_q \hat{\alpha}_l \hat{\pi}_{ql}}{2}$$

The connectivity between block q and l is the number of edges that connect a node from block q to l . Thus, we can formulate a random variable $E_{ql} = \sum_{i=1}^n \sum_{j < i}^n Z_{iq} Z_{jl} X_{ij}$ whose density $f(x_{ij}, z_{iq} = 1, z_{jl} = 1)$ is given as

$$f(x_{ij}, z_{iq} = 1, z_{jl} = 1) = P(Z_{iq} = 1)P(Z_{jl} = 1)P(X_{ij} = x_{ij}|Z_{iq} = 1, Z_{jl} = 1).$$

Next, the expected value of E_{ql} can be computed as

$$\begin{aligned}
\mathbb{E}(E_{ql}) &= \sum_{i=1}^n \sum_{j < i}^n z_{iq} P(Z_{iq} = 1) z_{jl} P(Z_{jl} = 1) x_{ij} f(x_{ij}|z_{iq} = 1, z_{jl} = 1) \\
&= \sum_{i=1}^n \sum_{j < i}^n \alpha_q \alpha_l x_{ij} \pi_{ql}^{x_{ij}} (1 - \pi_{ql})^{1-x_{ij}}
\end{aligned}$$

$$\begin{aligned}
&= \alpha_q \alpha_l \sum_{i=1}^n \sum_{j<i}^n x_{ij} \pi_{ql}^{x_{ij}} (1 - \pi_{ql})^{1-x_{ij}} \\
&= \alpha_q \alpha_l n(n-1)/2\pi_{ql}.
\end{aligned}$$

B.3 Coherent and Combinatory Blocks

Table B.1 shows the share of each anatomical regions in each blocks. Blocks which have 100% of their nodes from a specific anatomical regions are considered as Coherent. These blocks are marked by an asterisk. Blocks for which their nodes are from different anatomical regions are called Combinatory.

	Frontal Lobe	Temporal Lobe	Parietal Lobe	Occipital Lobe	Thalamus	Cerebellum	Limbic Lobe	Insular Lobe	Basal Ganglia
1	50	50							
2*						100			
3	70		15				15		
4*		100							
5	19	12	44					25	
6*				100					
7				55		45			
8*			100						
9				15	15	54	15		
10	50					7	29		14
11	33		66						
12	10	30					10	40	10
13	40		20				40		
14			33	44			22		
15				50			50		
16*		78	11					11	
17*			100						
18	6	27			6		53		6
19*							100		
20	27		72						
21*	100								
22		33	66						
23*	100								
24				45		55			
25					33				66
26*	100								
27*	100								
28		50		50					
29*	100								
30	50							50	
31			40				60		

Table B.1: List of blocks and the percentage that each anatomical region participated in them. Coherent blocks are marked by a black asterisk.

B.4 Projection of Estimated Het-SBM Block Assignments on the Brain Surface

Using Het-SBM on 13 healthy and 12 schizophrenic brain networks result in a 31-block model. As discussed in section 4.4.2, although blocks are obtained from functional activities of the subject without any prior knowledge of anatomy, spatial consistency can be seen across the brain as the regions known to be involved in specific functional tasks are grouped together. For instance, block 11 which represents the motor area of the brain.

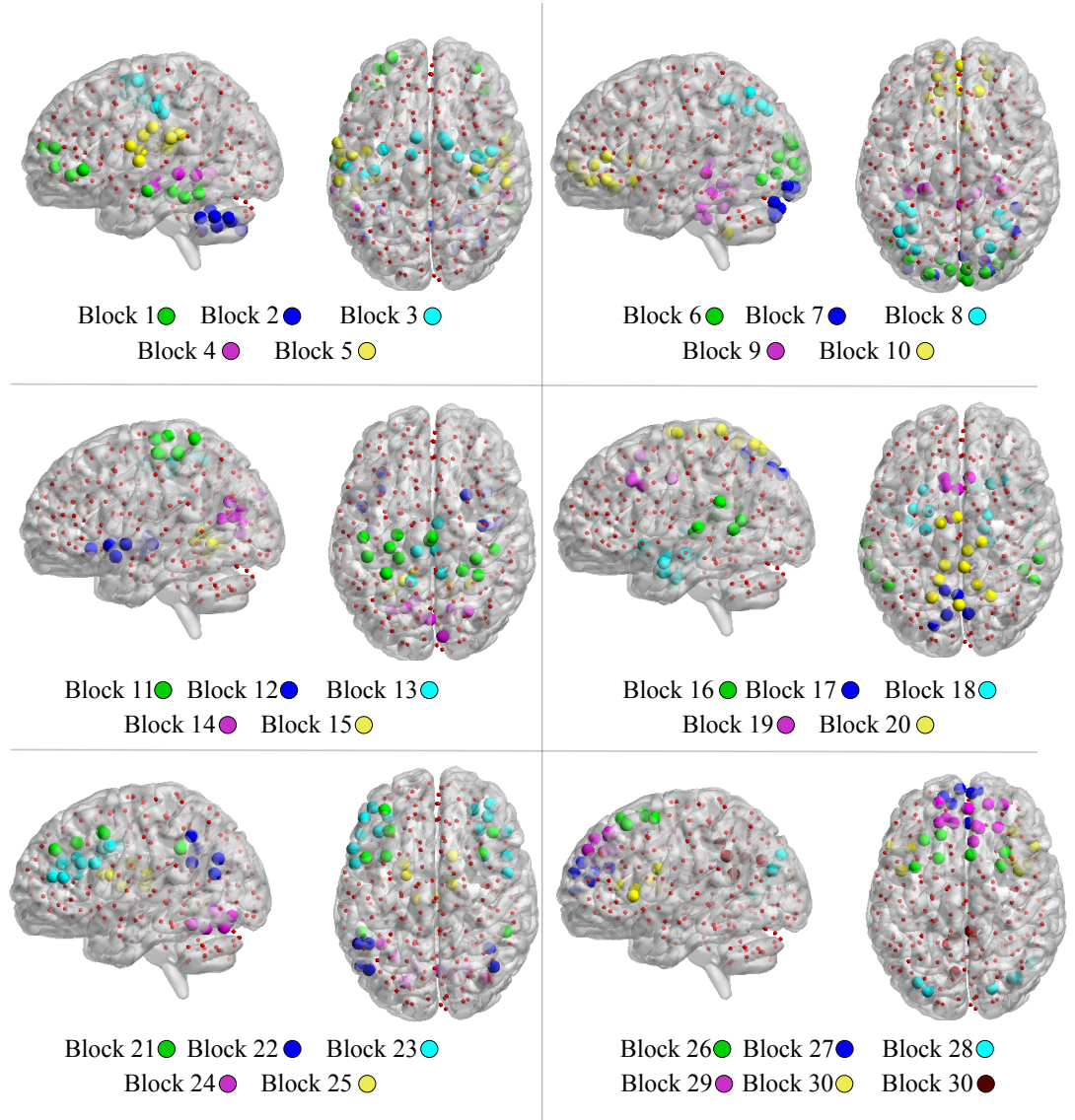


Figure B.1: Visualisation of block on MNI surface of the human brain. For the sake of visualisation, we show only 5 or 6 blocks per each brain surface.

B.5 Rich Block of C.Elegans Micro-scale Connectome

Further to validation of the RB coefficients presented in section 4.4.1, we use the micro-scale connectome of the C.Elegans to show how closely RB coefficients can estimate RC coefficients.

For this purpose we use a 282-node connectivity matrix of the C.Elegans described in Varshney *et al.* (2011). We discarded 3 isolated nodes to ensure that the network is connected (Towlson *et al.*, 2013). Following methods described in section 3.1.9, we determine the RC coefficients of the empirical network. In addition, we also estimate the block assignments using ERMM described in section 4.1.1. The optimal block assignments were found to have 9 blocks after 1k realisations. However, we do not merely suffice to the optimal model and explore the estimated block assignments for range of different block numbers $R=\{9, 20, 40, 70, 100\}$. For each member of set R , we calculate non-normalised RB coefficients, Υ_r^k , and plot the RB coefficients against their corresponding expected degree calculated using Eq 4.11.

Figure B.2 shows RB coefficients of different block numbers as well as empirical RC coefficients. The results suggest that as the number of blocks grows the RB coefficients show more similar results to the empirical RC coefficients. However, one should be careful when forcing the number of blocks beyond the optimal level as it may cause over-fitting.

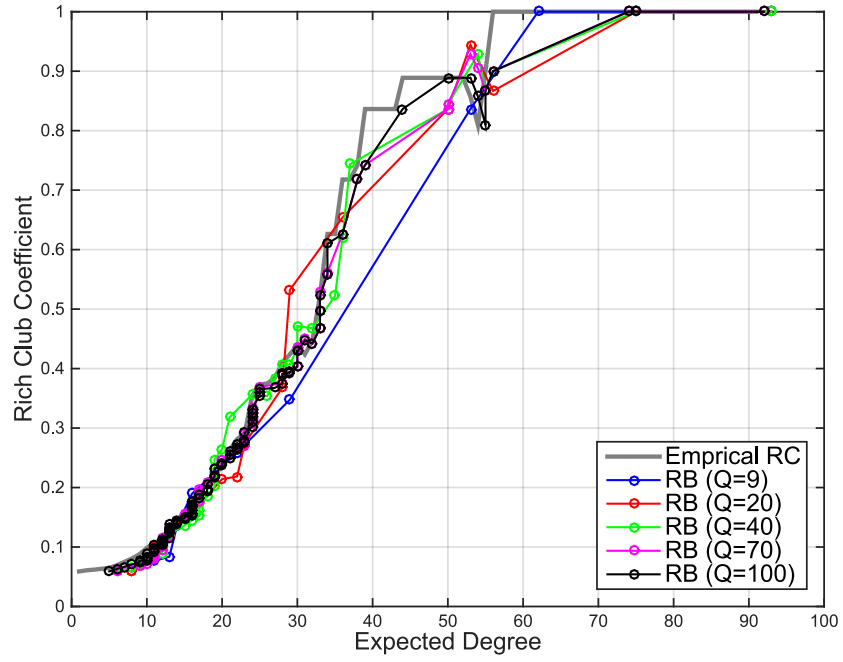


Figure B.2: Non-normalised RC (solid grey) and RB with optimal Q (marked blue) and fixed Q s (marked red, green, magenta and black)

B.6 Degree Exceptionality on Group-Level Het-SBM Parameters

Further to the discussion regarding detection of the hub blocks on individual level estimation of connectivity rate matrices of two groups, in section 4.4.4 and 4.1.4, we also show how the DE can be used to detect the hub blocks of the group-level.

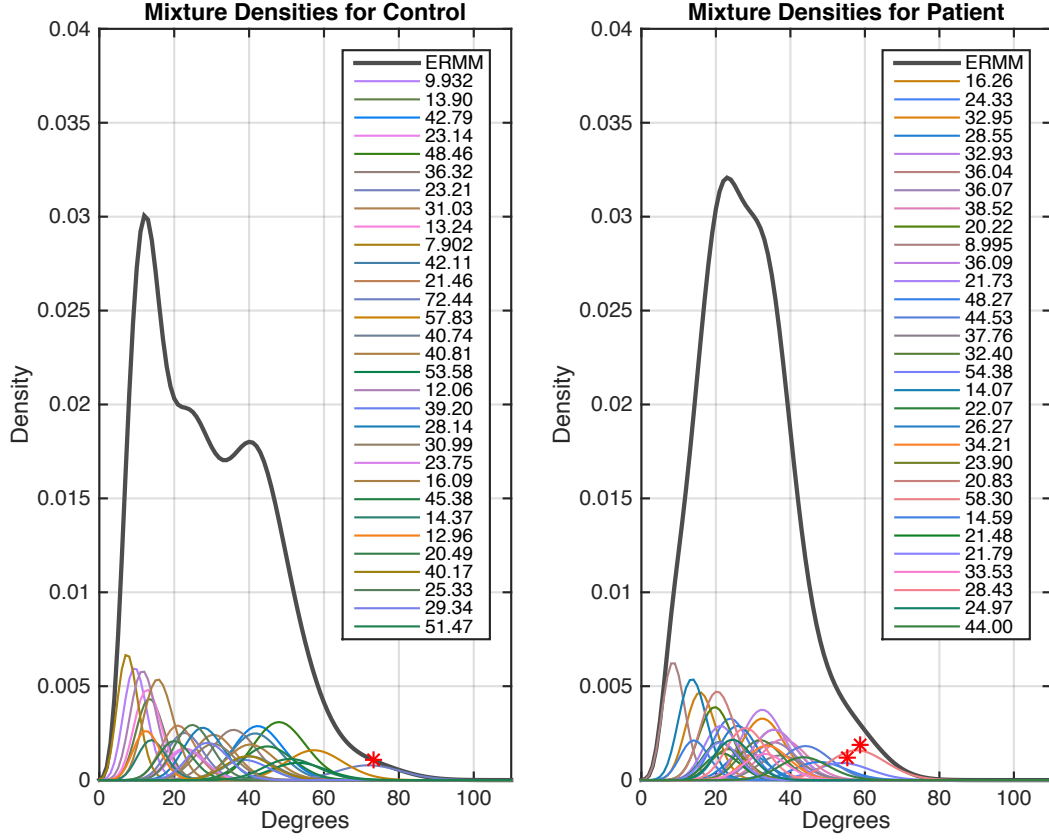


Figure B.3: Illustrates estimated PDF of the group-averaged connection probability matrix for controls (left) and patients (right). The red asterisk at the top of each Poisson component indicates degree exceptionality. Also the expected degree of each component involved in estimation of the degree distribution of each group is listed respective to their colour.

As discussed in section 4.4.2, we use Het-SBM to estimate the SBM parameters common between the two groups. The results of the Het-SBM suggest that the optimal number of blocks is 31. Using Eq 4.31, we detect blocks that are significantly different in terms of their expected degree compared to the rest of the blocks within each group. Figure B.3 shows the Poisson distribution of individual blocks for each group (Left: healthy, Right: schizophrenic) and the overall estimated degree distribution for each group. Results of DE suggest that the group-level blocks of healthy subjects merely suggest one hub block, whereas, in the schizophrenic group two blocks were suggested to be hub blocks. Using 4.11 we estimate the expected degree of hub blocks. Results suggest that the hub block in the healthy group has a degree of $\lambda_{13} = 72.44$, whereas, the hub blocks of schizophrenic subjects have $\lambda_{17} = 54.38$ and $\lambda_{24} = 58.30$. These results confirm the findings regarding the less connected blocks in schizophrenia at the individual level (see section 4.4.4).

Appendix C

Heterogeneous Bivariate Estimation of Sample Correlation Variance

C.1 Mean and Variance of Inner Product of Two Random Variables

Supposing that x and y are two random variable represented in the form of vectors, and \tilde{x} and \tilde{y} are deviations of each sample from their expectation, the mean and variance of the inner product between these two variables can be obtained as follows:

$$\begin{aligned} \langle x, y \rangle &= x'y \\ &= (\tilde{x} + \mu_x)(\tilde{y} + \mu_y) \\ &= \tilde{x}'\tilde{y} + 2\tilde{x}'\mu_y + 2\tilde{y}'\mu_x + \mu_x'\mu_y \end{aligned} \tag{C.1}$$

Noting that $E(\tilde{x}) = E(\tilde{y}) = 0$, the expected value of the inner product can be obtained as:

$$\begin{aligned} E \langle x, y \rangle &= E(\tilde{x}'\tilde{y}) + \mu_x'\mu_y \\ &= \text{Tr}(\Sigma_{xy}) + \mu_x'\mu_y \end{aligned} \tag{C.2}$$

And the variance as:

$$\begin{aligned} \text{var} \langle x, y \rangle &= E \langle x, y \rangle - [\langle x, y \rangle]^2 \\ &= E(\tilde{x}y)^2 - (\mu_x'\mu_y)^2 - 2\mu_x'\mu_y \text{Tr}(\Sigma_{xy}) - (\text{Tr}\Sigma_{xy})^2 \end{aligned} \tag{C.3}$$

Where $E(\tilde{x}y)$ can be obtained as:

$$\begin{aligned}
E(\tilde{x}y) &= E[(\tilde{x} + \mu_x)'(\tilde{y} + \mu_y)]^2 \\
&= E[\tilde{x}'\tilde{y}\tilde{y}'\tilde{x} + \mu_x'\tilde{y}\tilde{y}'\mu_x + \mu_y'\tilde{x}\tilde{x}'\mu_y + \\
&\quad (\mu_x'\mu_y)^2 + 2\mu_x'\tilde{y}\tilde{y}'\tilde{x} + 2\mu_y'\tilde{x}\tilde{x}'\tilde{y} + \\
&\quad 2\mu_x'\mu_y\tilde{x}\tilde{y} + 2\mu_y'\tilde{y}\tilde{x}\mu_y + 2\mu_y'\tilde{x}\mu_x'\mu_y + 2\mu_x'\tilde{y}\mu_x'\mu_y]
\end{aligned} \tag{C.4}$$

By remembering that $E(\tilde{y}\tilde{y}') = \Sigma_{yy}$, $E(\tilde{x}\tilde{x}') = \Sigma_{xx}$, $E(\tilde{x}'\tilde{y}) = \text{Tr}(\Sigma_{xy})$ and $E(\tilde{x}) = E(\tilde{y}) = 0$, we can rewrite Eq C.4 as:

$$\begin{aligned}
E(xy)^2 &= E(\tilde{x}'\tilde{y}\tilde{y}'\tilde{x}) + \mu_x'\Sigma_{yy}\mu_x + \mu_y'\Sigma_{xx}'\mu_y + \\
&\quad 2\mu_x'\Sigma_{xy}\mu_y + (\mu_x'\mu_y)^2 + 2\mu_x'E(\tilde{y}\tilde{y}'\tilde{x}) + \\
&\quad 2\mu_y'E(\tilde{x}\tilde{x}'\tilde{y}) + 2\mu_x'\mu_y\text{Tr}(\Sigma_{xy})
\end{aligned} \tag{C.5}$$

We note that

$$\mu'\Sigma^T\mu = \mu_x'\Sigma_{yy}\mu_x + \mu_y'\Sigma_{xx}'\mu_y + 2\mu_x'\Sigma_{xy}\mu_y \tag{C.6}$$

Hence,

$$\text{var} < x, y > = \mu'\Sigma^T\mu + 2\mu_x'E(\tilde{y}\tilde{y}'\tilde{x}) + 2\mu_y'E(\tilde{x}\tilde{x}'\tilde{y}) + E(\tilde{x}'\tilde{y}\tilde{y}'\tilde{x}) - \text{Tr}(\Sigma_{xy})^2 \tag{C.7}$$

C.2 Unibased-HetBiv vs Bartlett's HetBiv

In this section, we aim to show the differences between versions of the HetBiv-corrections by comparing their distribution shapes (via Kolmogorov Smirnov Statistics) and their false-positive rate (via tail analysis). The results which were presented later were obtained from inter-subject correlation for validating the methods, discussed in section 5.1.9. It is worth mentioning that due to computational infeasibilities, we drastically reduce the number of reaslisations which may introduce a difference in the KS statistics of the results presented in Figure C.1 and figure 5.3 for HetBiv.

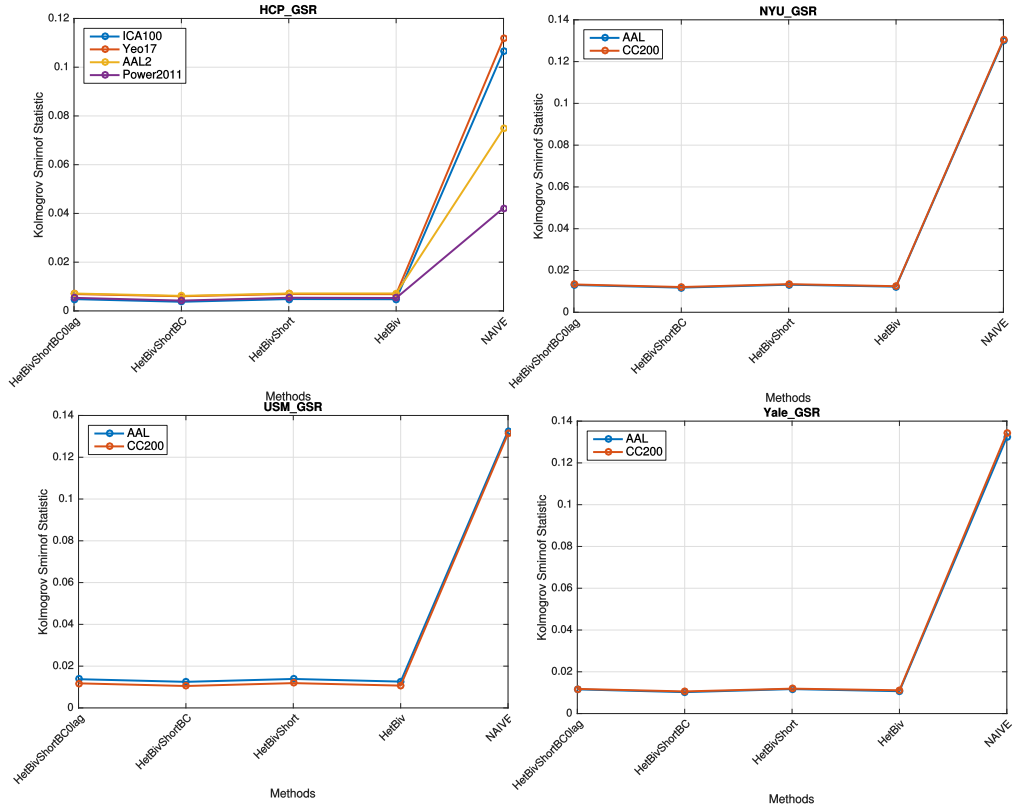


Figure C.1: KS Statistics of HetBiv methods. Each subplot represents a cohort: HCP (Top Left), NYU (Top Right), USM (Bottom Left) and Yale (Bottom Right).

Figure C.1 shows the KS statistics of the inter-subject correlations. The results suggest that different versions of the HetBiv-correction perform similarly, however, the Unbiased Short HetBiv slightly outperforms the other methods. It is worth noting that the KS results for individual parcellation schemes demonstrate a consistency across correction methods.

Figure C.2 shows analysis of tails, as discussed in section 5.1.9, across different versions of the HetBiv-correction. All versions of the HetBiv-correction could successfully adjust the tail of the distribution to avoid any false-positives in the statistical thresholding.

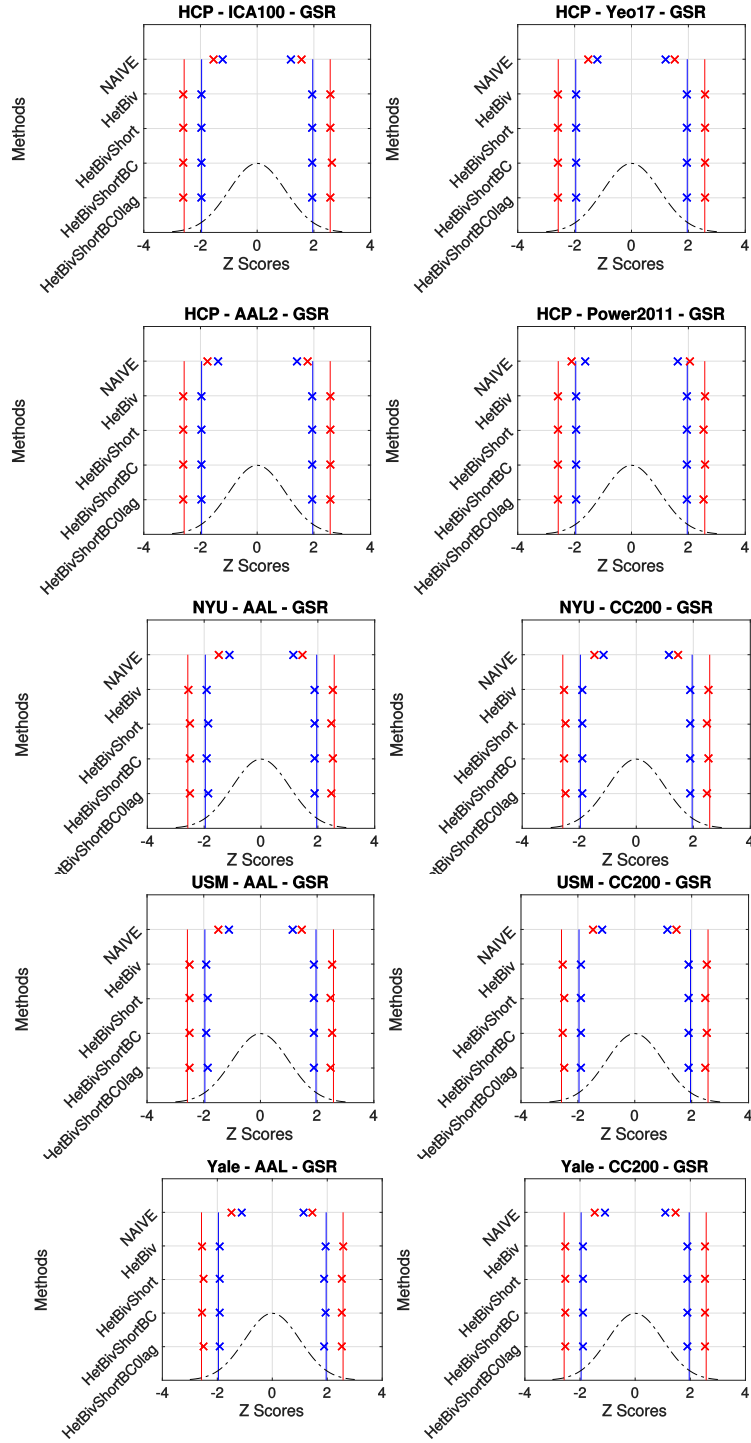


Figure C.2: False-positive rates of the HetBiv methods. Blue vertical lines represent the 5% α -level and red vertical lines represent 2.5% α -level.

C.3 Effect of Curbing in HetBiv Corrections

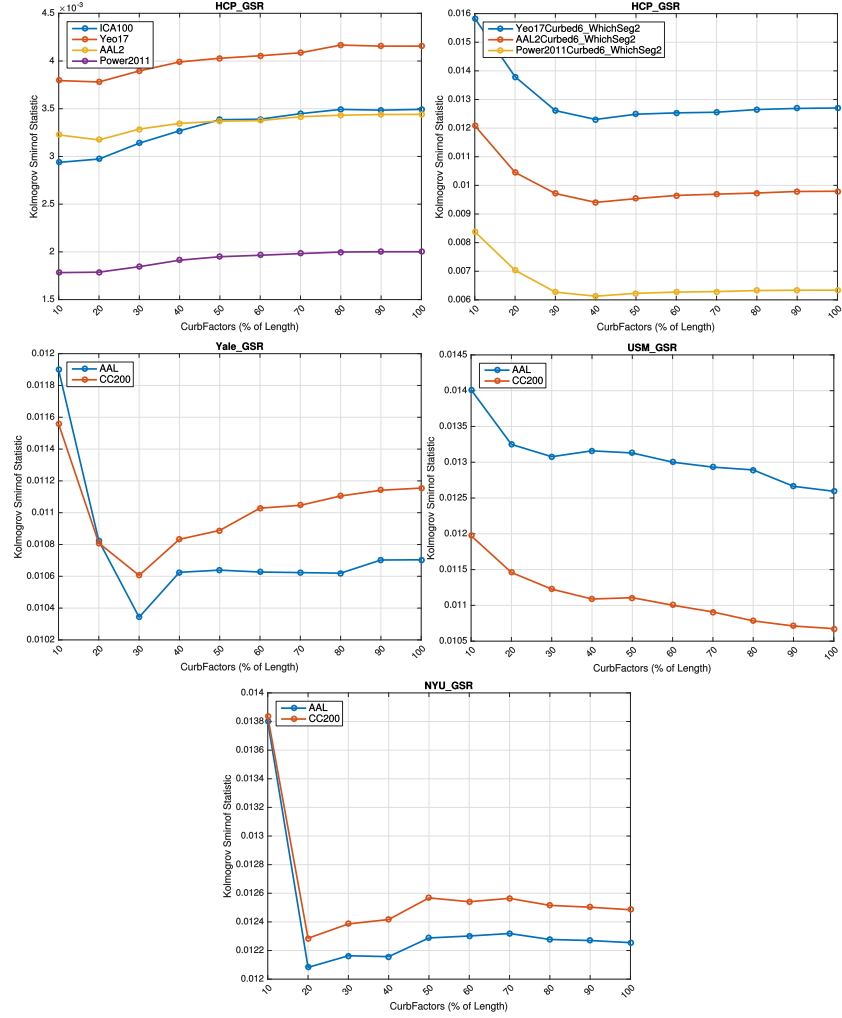


Figure C.3: KS statistics of each curbing coefficient across different acquisition sites and parcellation schemes.

In this section, we investigate the effect of choosing different curbing strategies for the length of the autocorrelation functions suggested in section 5.1.6. The curbing range starts from 10% and ends at 100% by 10% increments. The percentages represent the proportion of the original temporal lags that were involved in the calculation of the HetBiv-correction (Eq 5.28). In other words, the 100% curbing means that the correction method has used all available temporal lags which is equivalent to what is discussed in Eq 5.28. Further to conventional atlases that we discussed in section 5.2, we examined the curbing factors on a shortened versions of the HCP time series. For the shortened versions, we first divided each

HCP BOLD signal into six segments (each segments has 200 data-points), and used the second segment to ensure that the machine has reached the equilibrium state. The results presented here were obtained from the inter-subject validation technique.

Figure C.3 shows that the effect of curbing the temporal lags of an autocorrelation function is not consistent across sites. However, for each site, the results demonstrate a consistent change across parcellation schemes. This suggests that the curbing is sensitive to acquisition and pre-processing choices. For example, in the HCP cohort, harsh curbings, such as 10%-60%, improve the KS statistics, while, in the remaining conditions the same boundary shows a deterioration in the quality of correction.

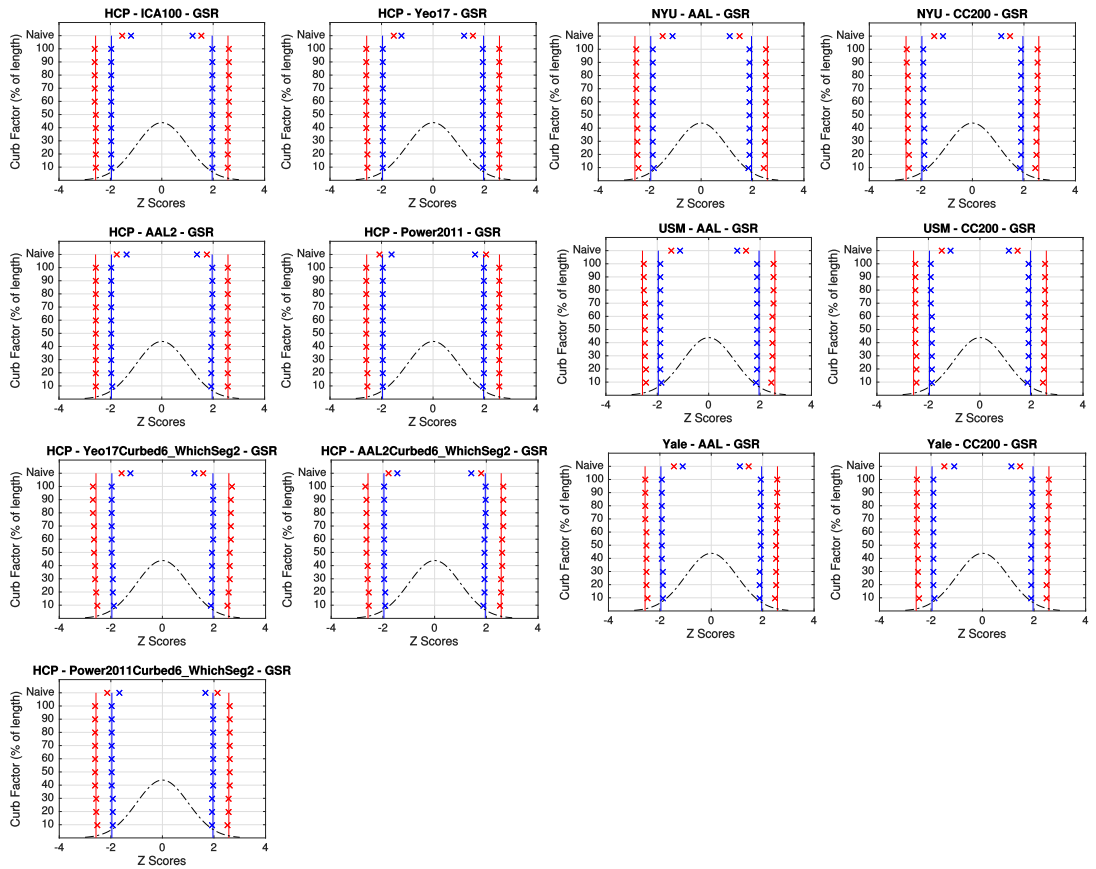


Figure C.4: False-positive rate analysis of each curbing coefficient across different acquisition sites and parcellation schemes.

C.4 Effect of Shrinking in HetBiv Corrections

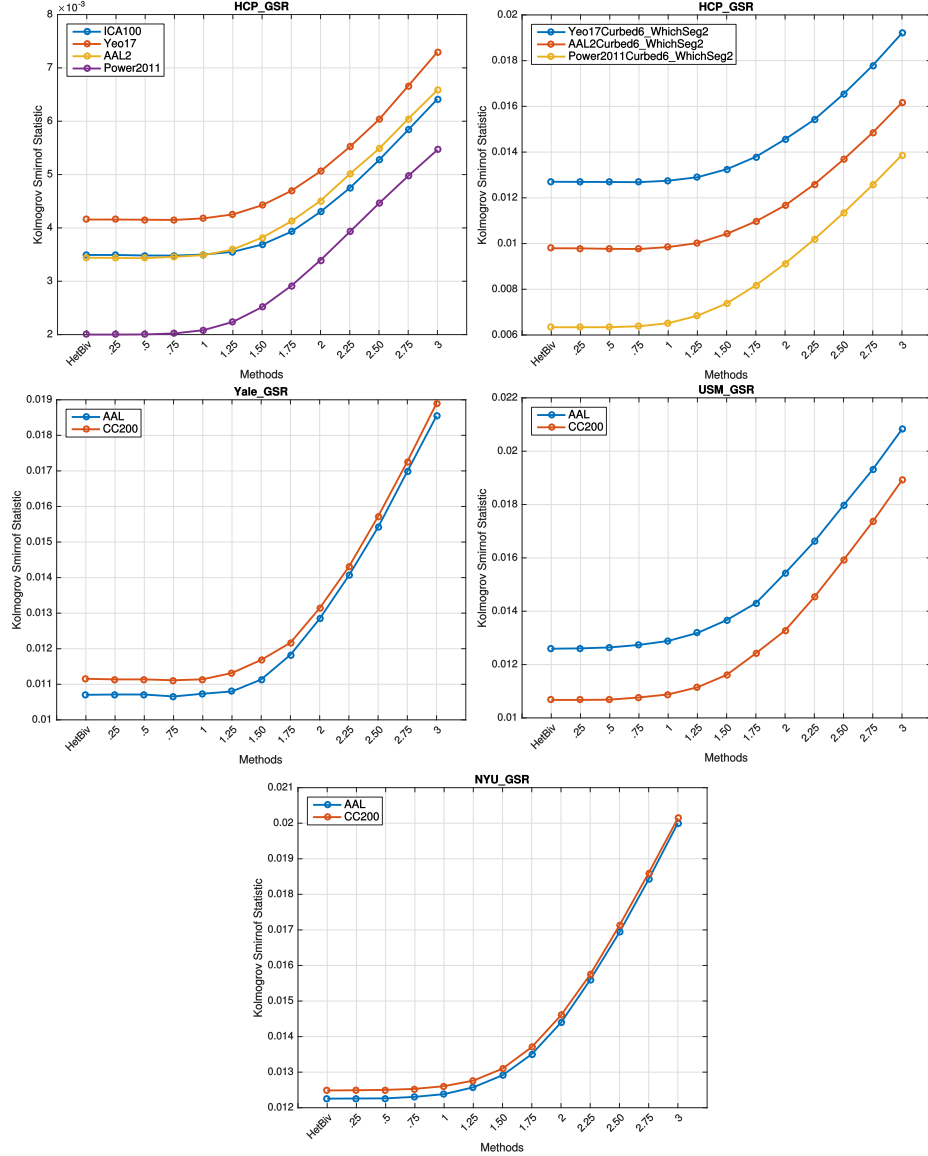


Figure C.5: KS statistics of each shrinking coefficient across different acquisition sites and parcellation schemes.

In this section, we show how the shrinking strategy effects the result of HetBiv correction. Shrinking is a coefficient of the number of standard deviations that are used to determine the boundaries. This begins from 0 and approaches 3 with steps of 0.25. The shrinking coefficient of 0 is equivalent to the original HetBiv correction described in Eq 5.28. Similar to section C.3, further to conventional atlases that were discussed in section 5.2, we examined

the curbing factors on shortened version of the HCP time series. The shortened versions, we first divided each HCP BOLD signal into six segments (each segment has 200 data-points), and used the second segment to ensure that the machine has reached the equilibrium state. The results presented here were obtained from inter-subject validation technique. The results presented here were obtained from the inter-subject correlation discussed in section 5.1.9.

Figure C.5 shows the difference between the distribution shapes which suggests a consistent change across sites and parcellation schemes over the shrinking coefficients, however, the results suggest that the shrinking coefficients do not introduce any improvement in KS statistics. However, for certain coefficients the result of the HetBiv corrections start to get worse as the KS statistic grows.

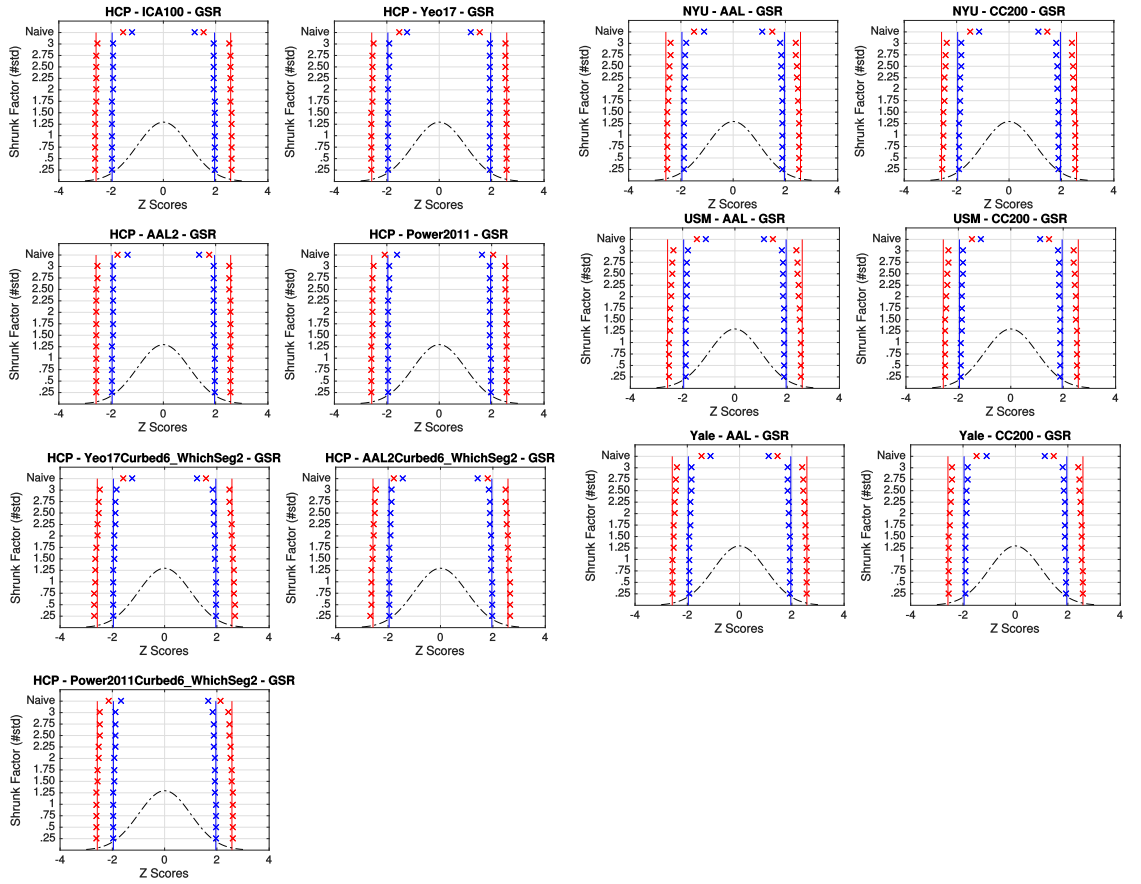


Figure C.6: False-positive rate of each shrinking coefficient across different acquisition sites and parcellation schemes.

Figure C.6 shows false-positive rates for each method across different cohorts and parcellation schemes. The results suggest that in long time series (such as the HCP cohort),

the shrinking coefficient neither improves nor worsens the results. However, for shorter time series from the NYU, Yale and USM cohorts, the growth in the coefficient is associated with the growth in the false-positive rate. Conversely, for curbed-time series of the HCP cohort, the result suggests that the false-positives are best controlled in coefficient 2.

Appendix D

List of Publications

D.1 Chapter 3

Abstracts:

- **Afyouni S**, Hale JR, Mayhew SD, Bagshaw AP, Arvanitis TN. Inter-network and Intra-network analysis of resting-state human brain. Poster presented at: Organisation for Human Brain Mapping: 20th Annual Meeting. 2014 June 8-12; Hamburg, Germany.
- **Afyouni S**, Hale JR, Mayhew SD, Bagshaw AP, and Arvanitis TN. Hub identification in dynamic resting-state functional connectivity of the default mode network. Poster presented at: Joint Annual Meeting ISMRM-ESMRMB 2014: SMRT 23rd Annual Meeting. 2014 May 10-16; Milan, Italy.

Journal paper under review and re-submission:

- Wilson R.S., Mayhew S.D., Hale J.R., Rollings D.T., **Afyouni S.**, Goldstone A., Khalsa S., Arvanitis T.N., Bagshaw A.P. The sleeping brains intrinsic network connectivity is changed by sleep deprivation, *Neuroimage*, (September 2015)

D.2 Chapter 4

Abstracts:

- **Afyouni S**, Pavlovic, D. M., Towlson, E. K., Petra, E. V. Nichols T. Bullmore E., Understanding Rich Club with a Stochastic Block Model: A Connectome Study of Schizophrenia. Poster presented at: Complex Networks: from Theory to Interdisciplinary Applications. 2016 July 10-13; Marseilles, France.

- **Afyouni S.**, Pavlovic D.M., Towlson E.K., Petra E.V., Arvanitis T.N., Bullmore E.T., Nichols T.E., Understanding Rich Club with a Stochastic Block Model: A Connectome Study of Resting-State Schizophrenia
- **Afyouni S.**, Pavlovic, D. M., Towlson, E. K., Petra, E. V. Nichols T. Bullmore E., Rich Block: Stochastic Understanding of Rich Club. Poster presented at: Organisation for Human Brain Mapping: 21th Annual Meeting. 2015 June 16-21; Hawaii, US.
- Pavlovic, D. M., Towlson, E. K., Petra, E. V. **Afyouni S.** Nichols T. Bullmore E., Stochastic Blockmodelling and Inference in Multi-Subject Networks, Poster presented at: Organisation for Human Brain Mapping: 21th Annual Meeting. 2015 June 16-21; Hawaii, US.

Journal paper under review and re-submission:

- **Afyouni S.**, Pavlovic D.M., Towlson E.K., Petra E.V., Arvanitis T.N., Bullmore E.T., Nichols T.E., Understanding Rich Club with a Stochastic Block Model: A Connectome Study of Resting-State Schizophrenia.
- Pavlovic, D.M., Towlson E.K., Petra E.V. **Afyouni S.**, Nichols T.E., Bullmore E.T., Stochastic Blockmodelling and Inference in Multi-Subject Networks, *Computational Statistics and Data Analysis*, (February 2015).

D.3 Chapter 5

Abstracts:

- **Afyouni S** and Nichols T E, Heritability of Global Architectural Features of the Functional Connectome of the Human Brain, Poster presented at: Organisation for Human Brain Mapping: 22nd Annual Meeting. 2016 June 25-30; Geneva, Switzerland.

Journal paper under preparation:

- **Afyouni S.** & Nichols T.E., Impact of Autocorrelation in Functional Human Connectomics .

Appendix E

Codes

The codes for this thesis are available in a separate DVD. The address to directories for each chapter is as follows:

- Chapter 3: `~ /SAfyouni/Schiz`
- Chapter 4: `~ /SAfyouni/Schiz/StochasticRB`
- Chapter 5: `~ /SAfyouni/HCP – Scripts/NetMats/Scripts/`

The codes are also available publicly through GitHub repositories as follows:

- Chapter 4: <https://github.com/asoroosh/HetBiv>
- Chapter 5: <https://github.com/asoroosh/RB>

By the time of submitting this thesis, the papers for chapter 4 and 5 are still in the review process, so the GitHub repositories are private. However, by requesting access to s.afyouni@warwick.ac.uk access can be granted.

Bibliography

- Achard, S. and Bullmore, E. (2007). Efficiency and cost of economical brain functional networks. *PLoS Comput Biol*, **3**(2), e17.
- Achard, S., Salvador, R., Whitcher, B., Suckling, J., and Bullmore, E. (2006). A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *The Journal of neuroscience*, **26**(1), 63–72.
- Achard, S., Delon-martin, C., Vértés, P. E., Renard, F., Schenck, M., and Schneider, F. (2012). Hubs of brain functional networks are radically reorganized in comatose patients.
- Adamic, L. A. and Glance, N. (2005). The political blogosphere and the 2004 us election: divided they blog. In *Proceedings of the 3rd international workshop on Link discovery*, pages 36–43. ACM.
- Albert, R. and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of modern physics*, **74**(1), 47.
- Alexander-Bloch, A., Giedd, J. N., and Bullmore, E. (2013). Imaging structural co-variance between human brain regions. *Nature Reviews Neuroscience*, **14**(5), 322–336.
- Alexander-Bloch, A. F., Gogtay, N., Meunier, D., Birn, R., Clasen, L., Lalonde, F., Lenroot, R., Giedd, J., and Bullmore, E. T. (2010). Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia. *Frontiers in systems neuroscience*, **4**, 147.
- Anderson, J. S., Druzgal, T. J., Froehlich, A., DuBray, M. B., Lange, N., Alexander, A. L., Abildskov, T., Nielsen, J. A., Cariello, A. N., Cooperrider, J. R., *et al.* (2010). Decreased interhemispheric functional connectivity in autism. *Cerebral cortex*, page bhq190.
- Anderson, O. D. (1983). *Time Series Analysis, Theory and Practice*, volume 7. North-Holland.

- Arbabshirani, M. R., Damaraju, E., Phlypo, R., Plis, S., Allen, E., Ma, S., Mathalon, D., Preda, A., Vaidya, J. G., Adali, T., *et al.* (2014). Impact of autocorrelation on functional connectivity. *Neuroimage*, **102**, 294–308.
- Ashburner, J., Friston, K. J., and Frackowiak, R. (2003). *Rigid body registration*. San Diego, CA: Academic Press.
- Ball, G., Aljabar, P., Zebari, S., Tusor, N., Arichi, T., Merchant, N., Robinson, E. C., Ogundipe, E., Rueckert, D., Edwards, A. D., *et al.* (2014). Rich-club organization of the newborn human brain. *Proceedings of the National Academy of Sciences*, **111**(20), 7456–7461.
- Barabási, A.-L. *et al.* (2009). Scale-free networks: a decade and beyond. *science*, **325**(5939), 412.
- Bartlett, M. (1935). Some aspects of the time-correlation problem in regard to tests of significance. *Journal of the Royal Statistical Society*, **98**(3), 536–543.
- Bartlett, M. S. (1946). On the theoretical specification and sampling properties of auto-correlated time-series. *Supplement to the Journal of the Royal Statistical Society*, **8**(1), 27–41.
- Bassett, D. S. and Bullmore, E. (2006). Small-world brain networks. *The neuroscientist*, **12**(6), 512–523.
- Bassett, D. S., Bullmore, E., Verchinski, B. A., Mattay, V. S., Weinberger, D. R., and Meyer-Lindenberg, A. (2008). Hierarchical organization of human cortical networks in health and schizophrenia. *The Journal of Neuroscience*, **28**(37), 9239–9248.
- Bassett, D. S., Bullmore, E. T., Meyer-Lindenberg, A., Apud, J. A., Weinberger, D. R., and Coppola, R. (2009). Cognitive fitness of cost-efficient brain functional networks. *Proceedings of the National Academy of Sciences*, **106**(28), 11747–11752.
- Bassett, D. S., Greenfield, D. L., Meyer-Lindenberg, A., Weinberger, D. R., Moore, S. W., and Bullmore, E. T. (2010). Efficient physical embedding of topologically complex information processing networks in brains and computer circuits. *PLoS Comput Biol*, **6**(4), e1000748.
- Bassett, D. S., Wymbs, N. F., Porter, M. A., Mucha, P. J., Carlson, J. M., and Grafton, S. T. (2011). Dynamic reconfiguration of human brain networks during learning. *Proceedings of the National Academy of Sciences*, **108**(18), 7641–7646.

- Becerril, K. E., Repovs, G., and Barch, D. M. (2011). Error processing network dynamics in schizophrenia. *Neuroimage*, **54**(2), 1495–1505.
- Beckmann, C. F., Mackay, C. E., Filippini, N., and Smith, S. M. (2009). Group comparison of resting-state fmri data using multi-subject ica and dual regression. *Neuroimage*, **47**(Suppl 1), S148.
- Behzadi, Y., Restom, K., Liao, J., and Liu, T. T. (2007). A component based noise correction method (compcor) for bold and perfusion based fmri. *Neuroimage*, **37**(1), 90–101.
- Bennett, C. M., Wolford, G. L., and Miller, M. B. (2009). The principled control of false positives in neuroimaging. *Social cognitive and affective neuroscience*, **4**(4), 417–422.
- Bernard, H. R., Killworth, P. D., Evans, M. J., McCarty, C., and Shelley, G. A. (1988). Studying social relations cross-culturally. *Ethnology*, **27**(2), 155–179.
- Bernard, H. R., Johnsen, E. C., Killworth, P. D., and Robinson, S. (1991). Estimating the size of an average personal network and of an event subpopulation: Some empirical results. *Social science research*, **20**(2), 109–121.
- Bernardo, J., Bayarri, M., Berger, J., Dawid, A., Heckerman, D., Smith, A., West, M., *et al.* (2003). The variational bayesian em algorithm for incomplete data: with application to scoring graphical model structures. *Bayesian statistics*, **7**, 453–464.
- Betzal, R. F., Griffa, A., Avena-Koenigsberger, A., Goñi, J., Thiran, J.-P., Hagmann, P., and Sporns, O. (2013). Multi-scale community organization of the human structural connectome and its relationship with resting-state functional connectivity. *Network Science*, **1**(03), 353–373.
- Biernacki, C., Celeux, G., and Govaert, G. (2000). Assessing a mixture model for clustering with the integrated completed likelihood. *IEEE transactions on pattern analysis and machine intelligence*, **22**(7), 719–725.
- Birn, R. M., Diamond, J. B., Smith, M. A., and Bandettini, P. A. (2006). Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fmri. *Neuroimage*, **31**(4), 1536–1548.
- Biswal, B., Zerrin Yetkin, F., Haughton, V. M., and Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magnetic resonance in medicine*, **34**(4), 537–541.
- Bland, J. M. and Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The lancet*, **327**(8476), 307–310.

- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, **2008**(10), P10008.
- Bouttier, J., Di Francesco, P., and Guitter, E. (2003). Geodesic distance in planar graphs. *Nuclear Physics B*, **663**(3), 535–567.
- Bressler, S. L. and Ding, M. (1999). Coordination dynamics in large-scale cortical networks. In *Neural Networks, 1999. IJCNN'99. International Joint Conference on*, volume 1, pages 113–116. IEEE.
- Brown, G. G. and Rutemiller, H. C. (1977). Means and variances of stochastic vector products with applications to random linear models. *Management Science*, **24**(2), 210–216.
- Bullmore, E. and Sporns, O. (2012a). The economy of brain network organization. *Nature Reviews Neuroscience*, **13**(5), 336–349.
- Bullmore, E. and Sporns, O. (2012b). The economy of brain network organization. *Nature reviews. Neuroscience*, **13**(5), 336–49.
- Bullmore, E., Long, C., Suckling, J., Fadili, J., Calvert, G., Zelaya, F., Carpenter, T. A., and Brammer, M. (2001). Colored noise and computational inference in neurophysiological (fmri) time series analysis: resampling methods in time and wavelet domains. *Human brain mapping*, **12**(2), 61–78.
- Bullmore, E. T. and Bassett, D. S. (2011). Brain graphs: graphical models of the human brain connectome. *Annual review of clinical psychology*, **7**, 113–140.
- Bullmore, E. T., Suckling, J., Overmeyer, S., Rabe-Hesketh, S., Taylor, E., and Brammer, M. J. (1999). Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural mr images of the brain. *IEEE transactions on medical imaging*, **18**(1), 32–42.
- Buxhoeveden, D. P. and Casanova, M. F. (2002). The minicolumn hypothesis in neuroscience. *Brain*, **125**(5), 935–951.
- Carbonell, F., Bellec, P., and Shmuel, A. (2011). Global and system-specific resting-state fmri fluctuations are uncorrelated: principal component analysis reveals anti-correlated networks. *Brain connectivity*, **1**(6), 496–510.

- Chabernaude, C., Mennes, M., Kelly, C., Nooner, K., Di Martino, A., Castellanos, F. X., and Milham, M. P. (2012). Dimensional brain-behavior relationships in children with attention-deficit/hyperactivity disorder. *Biological psychiatry*, **71**(5), 434–442.
- Chang, C. and Glover, G. H. (2009). Relationship between respiration, end-tidal co 2, and bold signals in resting-state fmri. *Neuroimage*, **47**(4), 1381–1393.
- Colizza, V., Flammini, A., Serrano, M. A., and Vespignani, A. (2006a). Detecting rich-club ordering in complex networks. *Nature physics*, **2**(2), 110–115.
- Colizza, V., Flammini, a., Serrano, M. a., and Vespignani, a. (2006b). Detecting rich-club ordering in complex networks. *Nature Physics*, **2**(2), 110–115.
- Collin, G., de Reus, M. A., Cahn, W., Pol, H. E. H., Kahn, R. S., and van den Heuvel, M. P. (2013). Disturbed grey matter coupling in schizophrenia. *European Neuropsychopharmacology*, **23**(1), 46–54.
- Collin, G., Kahn, R. S., de Reus, M. A., Cahn, W., and van den Heuvel, M. P. (2014a). Impaired rich club connectivity in unaffected siblings of schizophrenia patients. *Schizophrenia bulletin*, **40**(2), 438–448.
- Collin, G., Kahn, R. S., de Reus, M. A., Cahn, W., and van den Heuvel, M. P. (2014b). Impaired Rich Club Connectivity in Unaffected Siblings of Schizophrenia Patients. *Schizophrenia Bulletin*, **40**(2), 438–448.
- Collins, D. L., Holmes, C. J., Peters, T. M., and Evans, A. C. (1995). Automatic 3-d model-based neuroanatomical segmentation. *Human brain mapping*, **3**(3), 190–208.
- Collins, R. (2012). What makes uk biobank special? *The Lancet*, **379**(9822), 1173–1174.
- Costa, L., Smith, J., Nichols, T., Cussens, J., Duff, E. P., Makin, T. R., *et al.* (2015). Searching multiregression dynamic models of resting-state fmri networks using integer programming. *Bayesian Analysis*, **10**(2), 441–478.
- Craddock, R. C., James, G. A., Holtzheimer, P. E., Hu, X. P., and Mayberg, H. S. (2012). A whole brain fmri atlas generated via spatially constrained spectral clustering. *Human brain mapping*, **33**(8), 1914–1928.
- Craddock, R. C., Jbabdi, S., Yan, C.-G., Vogelstein, J. T., Castellanos, F. X., Di Martino, A., Kelly, C., Heberlein, K., Colcombe, S., and Milham, M. P. (2013). Imaging human connectomes at the macroscale. *Nature methods*, **10**(6), 524–539.
- Craven, P. and Wellman, B. (1973). The network city. *Sociological inquiry*, **43**(3-4), 57–88.

- Crossley, N. A., Mechelli, A., Scott, J., Carletti, F., Fox, P. T., McGuire, P., and Bullmore, E. T. (2014). The hubs of the human connectome are generally implicated in the anatomy of brain disorders. *Brain*, **137**(8), 2382–2395.
- Daniel, W. W. *et al.* (1990). Applied nonparametric statistics.
- Danon, L., Diaz-Guilera, A., Duch, J., and Arenas, A. (2005). Comparing community structure identification. *Journal of Statistical Mechanics: Theory and Experiment*, **2005**(09), P09008.
- Daudin, J.-j., Lacroix, V., Picard, F., Sagot, M.-f., Ina, U. M. R., Inra, P.-g. E., and Ssb, G. (2004). Uncovering structure in biological networks. (2003).
- Daudin, J.-J., Picard, F., and Robin, S. (2007). A mixture model for random graphs. *Statistics and Computing*, **18**(2), 173–183.
- de Reus, M. A. and van den Heuvel, M. P. (2014). Simulated rich club lesioning in brain networks: a scaffold for communication and integration? *Frontiers in human neuroscience*, **8**.
- Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., *et al.* (2006). An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. *Neuroimage*, **31**(3), 968–980.
- Di Martino, A., Yan, C.-G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., Anderson, J. S., Assaf, M., Bookheimer, S. Y., Dapretto, M., *et al.* (2014). The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular psychiatry*, **19**(6), 659–667.
- Dobson, I., Carreras, B. A., Lynch, V. E., and Newman, D. E. (2007). Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, **17**(2), 026103.
- Drakesmith, M., Caeyenberghs, K., Dutt, A., Zammit, S., Evans, C. J., Reichenberg, A., Lewis, G., David, A. S., and Jones, D. K. (2015). Schizophrenia-like topological changes in the structural connectome of individuals with subclinical psychotic experiences. *Human brain mapping*.
- Dubois, J. and Adolphs, R. (2016). Building a science of individual differences from fmri. *Trends in cognitive sciences*, **20**(6), 425–443.

- Edelman, G. M. (1987). *Neural Darwinism: The theory of neuronal group selection*. Basic Books.
- Edelman, G. M. (1993). Neural darwinism: selection and reentrant signaling in higher brain function. *Neuron*, **10**(2), 115–125.
- Eickhoff, S. B., Paus, T., Caspers, S., Grosbras, M.-H., Evans, A. C., Zilles, K., and Amunts, K. (2007). Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. *Neuroimage*, **36**(3), 511–521.
- Erdős, P. and Rényi, A. (1959). On random graphs, i. *Publicationes Mathematicae (Debrecen)*, **6**, 290–297.
- Euler, L. (1741). Solutio problematis ad geometriam situs pertinentis. *Commentarii academiae scientiarum Petropolitanae*, **8**, 128–140.
- Fararo, T. J. and Sunshine, M. H. (1964). *A study of a biased friendship net*. Youth Development Center, Syracuse University.
- Félix, M.-A. and Braendle, C. (2010). The natural history of *Caenorhabditis elegans*. *Current Biology*, **20**(22), R965–R969.
- Felleman, D. J. and Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex*, **1**(1), 1–47.
- Ferrarini, L., Veer, I. M., Baerends, E., van Tol, M.-J., Renken, R. J., van der Wee, N. J., Veltman, D., Aleman, A., Zitman, F. G., Penninx, B. W., *et al.* (2009). Hierarchical functional modularity in the resting-state human brain. *Human brain mapping*, **30**(7), 2220–2231.
- Fienberg, S. E. and Wasserman, S. (1981). Categorical data analysis of single sociometric relations. *Sociological methodology*, **12**, 156–192.
- Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, **80**(1), 27–38.
- Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, **10**(4), 507–521.
- Fornito, A., Zalesky, A., Bassett, D. S., Meunier, D., Ellison-Wright, I., Yücel, M., Wood, S. J., Shaw, K., O’Connor, J., Nertney, D., *et al.* (2011). Genetic influences on cost-efficient organization of human cortical functional networks. *The Journal of Neuroscience*, **31**(9), 3261–3270.

- Fornito, A., Zalesky, A., and Breakspear, M. (2013). Graph analysis of the human connectome: promise, progress, and pitfalls. *Neuroimage*, **80**, 426–444.
- Fornito, A., Zalesky, A., and Bullmore, E. (2016). *Fundamentals of Brain Network Analysis*. Academic Press.
- Fortunato, S. (2010). Community detection in graphs. *Physics reports*, **486**(3), 75–174.
- Fortunato, S. and Barthelemy, M. (2007). Resolution limit in community detection. *Proceedings of the National Academy of Sciences*, **104**(1), 36–41.
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., and Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(27), 9673–9678.
- Fox, M. D., Zhang, D., Snyder, A. Z., and Raichle, M. E. (2009). The global signal and observed anticorrelated resting state brain networks. *Journal of neurophysiology*, **101**(6), 3270–3283.
- Friston, K., Josephs, O., Zarahn, E., Holmes, A., Rouquette, S., and Poline, J.-B. (2000). To smooth or not to smooth?: Bias and efficiency in fmri time-series analysis. *NeuroImage*, **12**(2), 196–208.
- Friston, K. J., Holmes, A. P., Poline, J., Grasby, P., Williams, S., Frackowiak, R. S., and Turner, R. (1995). Analysis of fmri time-series revisited. *Neuroimage*, **2**(1), 45–53.
- Fulcher, B. D. and Fornito, A. (2016). A transcriptional signature of hub connectivity in the mouse connectome. *Proceedings of the National Academy of Sciences*, **113**(5), 1435–1440.
- Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J. R., *et al.* (2013). The minimal preprocessing pipelines for the human connectome project. *Neuroimage*, **80**, 105–124.
- Golestani, A.-M. and Goodyear, B. G. (2011). Regions of interest for resting-state fmri analysis determined by inter-voxel cross-correlation. *Neuroimage*, **56**(1), 246–251.
- Goparaju, B., Rana, K. D., Calabro, F. J., and Vaina, L. M. (2014). A computational study of whole-brain connectivity in resting state and task fmri. *Medical science monitor: international medical journal of experimental and clinical research*, **20**, 1024.

- Grayson, D. S., Ray, S., Carpenter, S., Iyer, S., Dias, T. G. C., Stevens, C., Nigg, J. T., and Fair, D. A. (2014). Structural and functional rich club organization of the brain in children and adults. *PloS one*, **9**(2), e88297.
- Hagmann, P., Jonasson, L., Maeder, P., Thiran, J.-P., Wedeen, V. J., and Meuli, R. (2006). Understanding diffusion mr imaging techniques: from scalar diffusion-weighted imaging to diffusion tensor imaging and beyond 1. *Radiographics*, **26**(suppl_1), S205–S223.
- Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., and Sporns, O. (2008). Mapping the structural core of human cerebral cortex. *PLoS Biol*, **6**(7), e159.
- Hale, J. R., White, T. P., Mayhew, S. D., Wilson, R. S., Rollings, D. T., Khalsa, S., Arvanitis, T. N., and Bagshaw, A. P. (2016). Altered thalamocortical and intra-thalamic functional connectivity during light sleep compared with wake. *NeuroImage*, **125**, 657–667.
- Harnad, S. (1994). Why and how we are not zombies. *Journal of Consciousness Studies*, **1**(2), 164–167.
- Hayasaka, S. and Nichols, T. E. (2004). Combining voxel intensity and cluster extent with permutation test framework. *Neuroimage*, **23**(1), 54–63.
- Henson, R., Buechel, C., Josephs, O., and Friston, K. (1999). The slice-timing problem in event-related fmri. *NeuroImage*, **9**, 125.
- Hosseini, S. H. and Kesler, S. R. (2013). Influence of choice of null network on small-world parameters of structural correlation networks. *PloS one*, **8**(6), e67354.
- Humphries, M. D. and Gurney, K. (2008). Network small-world-ness: a quantitative method for determining canonical network equivalence. *PloS one*, **3**(4), e0002051.
- Humphries, M. D., Gurney, K., and Prescott, T. J. (2006). The brainstem reticular formation is a small-world, not scale-free, network. *Proceedings of the Royal Society of London B: Biological Sciences*, **273**(1585), 503–511.
- Hurwicz, L. (1944). Stochastic models of economic fluctuations. *Econometrica: Journal of the Econometric Society*, pages 114–124.
- Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., Della Penna, S., Duyn, J. H., Glover, G. H., Gonzalez-Castillo, J., *et al.* (2013). Dynamic functional connectivity: promise, issues, and interpretations. *Neuroimage*, **80**, 360–378.

- Jaakkola, T. S. (2001). 10 tutorial on variational approximation methods. *Advanced mean field methods: theory and practice*, page 129.
- Jenkins, G. and Reinsel, G. (1976). Time series analysis: forecasting and control.
- Johnston, J. M., Vaishnavi, S. N., Smyth, M. D., Zhang, D., He, B. J., Zempel, J. M., Shimony, J. S., Snyder, A. Z., and Raichle, M. E. (2008). Loss of resting interhemispheric functional connectivity after complete section of the corpus callosum. *The Journal of neuroscience*, **28**(25), 6453–6458.
- Jordan, M. I., Ghahramani, Z., Jaakkola, T. S., and Saul, L. K. (1999). Introduction to variational methods for graphical models. *Machine Learning*, **37**(2), 183–233.
- Kaiser, M. (2008). Mean clustering coefficients: the role of isolated nodes and leafs on clustering measures for small-world networks. *New Journal of Physics*, **10**(8), 083042.
- Kalos, M. H. and Whitlock, P. A. (2008). *Monte carlo methods*. John Wiley & Sons.
- Karrer, B. and Newman, M. E. (2011). Stochastic blockmodels and community structure in networks. *Physical Review E*, **83**(1), 016107.
- Kenyon, C. (1988). The nematode *caenorhabditis elegans*. *Science*, **240**(4858), 1448.
- Kivelä, M., Arenas, A., Barthélemy, M., Gleeson, J. P., Moreno, Y., and Porter, M. A. (2014). Multilayer networks. *Journal of complex networks*, **2**(3), 203–271.
- Kleinberg, J. M. (2000). Navigation in a small world. *Nature*, **406**(6798), 845–845.
- Klovdahl, A. S., Potterat, J. J., Woodhouse, D. E., Muth, J. B., Muth, S. Q., and Darrow, W. W. (1994). Social networks and infectious disease: The colorado springs study. *Social science & medicine*, **38**(1), 79–88.
- Krackhardt, D. (1987). Cognitive social structures. *Social networks*, **9**(2), 109–134.
- Krebs, V. E. (2002). Mapping networks of terrorist cells. *Connections*, **24**(3), 43–52.
- Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *The annals of mathematical statistics*, **22**(1), 79–86.
- Lancichinetti, A. and Fortunato, S. (2011). Limits of modularity maximization in community detection. *Physical review E*, **84**(6), 066122.
- Latora, V. and Marchiori, M. (2001). Efficient behavior of small-world networks. *Physical review letters*, **87**(19), 198701.

- Lauterbur, P. C. (1973). Image formation by induced local interactions: examples employing nuclear magnetic resonance.
- Lein, E. S., Hawrylycz, M. J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A. F., Boguski, M. S., Brockway, K. S., Byrnes, E. J., *et al.* (2007). Genome-wide atlas of gene expression in the adult mouse brain. *Nature*, **445**(7124), 168–176.
- Liu, Y., Liang, M., Zhou, Y., He, Y., Hao, Y., Song, M., Yu, C., Liu, H., Liu, Z., and Jiang, T. (2008). Disrupted small-world networks in schizophrenia. *Brain*, **131**(4), 945–961.
- Logothetis, N. K. and Wandell, B. A. (2004). Interpreting the bold signal. *Annu. Rev. Physiol.*, **66**, 735–769.
- Ludbrook, J. (2008). Analysis of 2×2 tables of frequencies: matching test to experimental design. *International journal of epidemiology*, **37**(6), 1430–1435.
- Lund, T. E., Madsen, K. H., Sidaros, K., Luo, W.-L., and Nichols, T. E. (2006). Non-white noise in fmri: does modelling have an impact? *Neuroimage*, **29**(1), 54–66.
- Lynall, M.-E., Bassett, D. S., Kerwin, R., McKenna, P. J., Kitzbichler, M., Muller, U., and Bullmore, E. (2010). Functional connectivity and brain networks in schizophrenia. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **30**(28), 9477–87.
- Mariadassou, M., Robin, S., and Vacher, C. (2010). Uncovering latent structure in valued graphs: A variational approach. *Annals of Applied Statistics*, **4**(2), 715–742.
- Maslov, S. and Sneppen, K. (2002). Specificity and stability in topology of protein networks. *Science (New York, N.Y.)*, **296**(5569), 910–3.
- McAuley, J. J., Costa, L. D. F., and Caetano, T. S. (2007). The rich-club phenomenon across complex network hierarchies. page 4.
- McLachlan, G. and Peel, D. (2004). *Finite mixture models*. John Wiley & Sons.
- Meunier, D., Lambiotte, R., Fornito, A., Ersche, K. D., and Bullmore, E. T. (2010). Hierarchical modularity in human brain functional networks. *Hierarchy and dynamics in neural networks*, **1**, 2.
- Meunier, D., Fonlupt, P., Saive, A.-L., Plailly, J., Ravel, N., and Royet, J.-P. (2014). Modular structure of functional networks in olfactory memory. *NeuroImage*, **95**, 264–275.

- Micheloyannis, S., Pachou, E., Stam, C. J., Breakspear, M., Bitsios, P., Vourkas, M., Erimaki, S., and Zervakis, M. (2006). Small-world networks and disturbed functional connectivity in schizophrenia. *Schizophrenia research*, **87**(1), 60–66.
- Milgram, S. (1967). The small world problem. *Psychology today*, **2**(1), 60–67.
- Milward, H. B. and Raab, J. (2002). Dark networks: the structure, operation, and performance of international drug, terror, and arms trafficking networks. In *International Conference on the Empirical Study of Governance, Management, and Performance, Barcelona, Spain*.
- Mišić, B. and Sporns, O. (2016). From regions to connections and networks: new bridges between brain and behavior. *Current opinion in neurobiology*, **40**, 1–7.
- Monechi, B., Servedio, V. D., and Loreto, V. (2015). Congestion transition in air traffic networks. *PloS one*, **10**(5), e0125546.
- Moreno, J. L. (1934). *Who shall survive*, volume 58. JSTOR.
- Mucha, P. J., Richardson, T., Macon, K., Porter, M. A., and Onnela, J.-P. (2010). Community structure in time-dependent, multiscale, and multiplex networks. *science*, **328**(5980), 876–878.
- Murphy, K., Birn, R. M., Handwerker, D. A., Jones, T. B., and Bandettini, P. A. (2009). The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? *Neuroimage*, **44**(3), 893–905.
- Newman, M. (2010). *Networks: an introduction*. Oxford University Press.
- Newman, M. E. (2001). Scientific collaboration networks. i. network construction and fundamental results. *Physical review E*, **64**(1), 016131.
- Newman, M. E. (2002). The structure and function of networks. *Computer Physics Communications*, **147**(1), 40–45.
- Newman, M. E. (2003a). Mixing patterns in networks. *Physical Review E*, **67**(2), 026126.
- Newman, M. E. (2005). A measure of betweenness centrality based on random walks. *Social networks*, **27**(1), 39–54.
- Newman, M. E. (2006). Modularity and community structure in networks. *Proceedings of the national academy of sciences*, **103**(23), 8577–8582.

- Newman, M. E. (2013). Spectral methods for community detection and graph partitioning. *Physical Review E*, **88**(4), 042822.
- Newman, M. E. J. (2003b). The Structure and Function of Complex Networks. *SIAM Review*, **45**(2), 167–256.
- Nichols, T. and Hayasaka, S. (2003). Controlling the familywise error rate in functional neuroimaging: a comparative review. *Statistical methods in medical research*, **12**(5), 419–446.
- Nowicki, K. and Snijders, T. A. B. (2007). Estimation and Prediction for Stochastic Block-structures. **96**(455), 1077–1087.
- Olhede, S. C. and Wolfe, P. J. (2014). Network histograms and universality of blockmodel approximation. *Proceedings of the National Academy of Sciences*, **111**(41), 14722–14727.
- Ottet, M.-C., Schaer, M., Debbané, M., Cammoun, L., Thiran, J.-P., and Eliez, S. (2013). Graph theory reveals dysconnected hubs in 22q11ds and altered nodal efficiency in patients with hallucinations.
- Patel, A. X. and Bullmore, E. T. (2015). A wavelet-based estimator of the degrees of freedom in denoised fmri time series for probabilistic testing of functional connectivity and brain graphs. *NeuroImage*.
- Patel, A. X., Kundu, P., Rubinov, M., Jones, P. S., Vértes, P. E., Ersche, K. D., Suckling, J., and Bullmore, E. T. (2014). A wavelet method for modeling and despiking motion artifacts from resting-state fmri time series. *Neuroimage*, **95**, 287–304.
- Pavlovic, D. M., Vértes, P. E., Bullmore, E. T., Schafer, W. R., and Nichols, T. E. (2014). Stochastic blockmodeling of the modules and core of the *Caenorhabditis elegans* connectome. *PloS one*, **9**(7), e97584.
- Pavlovic, D. M., Towlson, E. K., and Petra, E. V. (2015). Stochastic Blockmodelling and Inference in Multi-Subject Networks. (February).
- Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., and Nichols, T. E. (2011). *Statistical parametric mapping: the analysis of functional brain images*. Academic press.
- Percival, D. B. and Walden, A. T. (2006). *Wavelet methods for time series analysis*, volume 4. Cambridge University Press.

- Poline, J.-B., Breeze, J. L., Ghosh, S. S., Gorgolewski, K., Halchenko, Y. O., Hanke, M., Helmer, K. G., Marcus, D. S., Poldrack, R. A., Schwartz, Y., *et al.* (2012). Data sharing in neuroimaging research. *Frontiers in neuroinformatics*, **6**, 9.
- Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., Vogel, A. C., Laumann, T. O., Miezin, F. M., Schlaggar, B. L., *et al.* (2011). Functional network organization of the human brain. *Neuron*, **72**(4), 665–678.
- Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., and Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fmri. *Neuroimage*, **84**, 320–341.
- Raichle, M. E. and Snyder, A. Z. (2007). A default mode of brain function: a brief history of an evolving idea. *Neuroimage*, **37**(4), 1083–1090.
- Rivera-Alba, M., Vitaladevuni, S. N., Mishchenko, Y., Lu, Z., Takemura, S.-y., Scheffer, L., Meinertzhagen, I. A., Chklovskii, D. B., and de Polavieja, G. G. (2011). Wiring economy and volume exclusion determine neuronal placement in the drosophila brain. *Current Biology*, **21**(23), 2000–2005.
- Roche, R. and Commins, S. (2009). *Pioneering Studies in Cognitive Neuroscience*. McGraw-Hill Education (UK).
- Roebroeck, A., Formisano, E., and Goebel, R. (2005). Mapping directed influence over the brain using granger causality and fmri. *Neuroimage*, **25**(1), 230–242.
- Rolls, E. T., Joliot, M., and Tzourio-Mazoyer, N. (2015). Implementation of a new parcellation of the orbitofrontal cortex in the automated anatomical labeling atlas. *NeuroImage*, **122**, 1–5.
- Rubinov, M. and Bullmore, E. (2013). Schizophrenia and abnormal brain network hubs. *Dialogues Clin Neurosci*, **15**(3), 339–349.
- Rubinov, M. and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*, **52**(3), 1059–1069.
- Rubinov, M., Ypma, R. J., Watson, C., and Bullmore, E. T. (2015). Wiring cost and topological participation of the mouse brain connectome. *Proceedings of the National Academy of Sciences*, **112**(32), 10032–10037.
- Salimi-Khorshidi, G., Douaud, G., Beckmann, C. F., Glasser, M. F., Griffanti, L., and Smith, S. M. (2014). Automatic denoising of functional mri data: combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage*, **90**, 449–468.

- Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G. C., *et al.* (1998). The mini-international neuropsychiatric interview (mini): the development and validation of a structured diagnostic psychiatric interview for dsm-iv and icd-10. *Journal of clinical psychiatry*, **59**, 22–33.
- Sheline, Y. I., Barch, D. M., Price, J. L., Rundle, M. M., Vaishnavi, S. N., Snyder, A. Z., Mintun, M. A., Wang, S., Coalson, R. S., and Raichle, M. E. (2009). The default mode network and self-referential processes in depression. *Proceedings of the National Academy of Sciences*, **106**(6), 1942–1947.
- Sinclair, B., Hansell, N. K., Blokland, G. A., Martin, N. G., Thompson, P. M., Breakspear, M., de Zubicaray, G. I., Wright, M. J., and McMahon, K. L. (2015). Heritability of the network architecture of intrinsic brain functional connectivity. *Neuroimage*, **121**, 243–252.
- Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., Ramsey, J. D., and Woolrich, M. W. (2011a). Network modelling methods for FMRI. *NeuroImage*, **54**(2), 875–91.
- Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., Ramsey, J. D., and Woolrich, M. W. (2011b). Network modelling methods for fmri. *Neuroimage*, **54**(2), 875–891.
- Smith, S. M., Beckmann, C. F., Andersson, J., Auerbach, E. J., Bijsterbosch, J., Douaud, G., Duff, E., Feinberg, D. A., Griffanti, L., Harms, M. P., *et al.* (2013). Resting-state fmri in the human connectome project. *Neuroimage*, **80**, 144–168.
- Smith, S. M., Nichols, T. E., Vidaurre, D., Winkler, A. M., Behrens, T. E., Glasser, M. F., Ugurbil, K., Barch, D. M., Van Essen, D. C., and Miller, K. L. (2015). A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nature neuroscience*, **18**(11), 1565–1567.
- Solomonoff, R. and Rapoport, A. (1951). Connectivity of random nets. *The bulletin of mathematical biophysics*, **13**(2), 107–117.
- Spitzer, R. L. and Williams, J. B. (1980). Diagnostic and statistical manual of mental disorders. In *American Psychiatric Association*. Citeseer.
- Sporns, O. (2011a). The human connectome: a complex network. *Annals of the New York Academy of Sciences*, **1224**(1), 109–125.
- Sporns, O. (2011b). The non-random brain: efficiency, economy, and complex dynamics. *Frontiers in computational neuroscience*, **5**, 5.

- Sporns, O. (2013). Structure and function of complex brain networks. *Dialogues Clin Neurosci*, **15**(3), 247–262.
- Sporns, O. and Honey, C. J. (2006). Small worlds inside big brains. *Proceedings of the National Academy of Sciences*, **103**(51), 19219–19220.
- Sporns, O. and Zwi, J. D. (2004). The Small World of the Cerebral Cortex. *Cerebral Cortex*, pages 145–162.
- Sporns, O., Chialvo, D. R., Kaiser, M., and Hilgetag, C. C. (2004). Organization, development and function of complex brain networks. *Trends in cognitive sciences*, **8**(9), 418–425.
- Sporns, O., Tononi, G., and Kötter, R. (2005). The human connectome: a structural description of the human brain. *PLoS Comput Biol*, **1**(4), e42.
- Sporns, O., Honey, C. J., and Kötter, R. (2007). Identification and classification of hubs in brain networks. *PloS one*, **2**(10), e1049.
- Stam, C. J. (2004). Functional connectivity patterns of human magnetoencephalographic recordings: a small-worldnetwork? *Neuroscience letters*, **355**(1), 25–28.
- Stam, C. J. (2014). Modern network science of neurological disorders. *Nature Reviews Neuroscience*, **15**(10), 683–695.
- Stam, C. J. and Reijneveld, J. C. (2007). Graph theoretical analysis of complex networks in the brain. *Nonlinear biomedical physics*, **1**(1), 1.
- Stepanyants, A., Hof, P. R., and Chklovskii, D. B. (2002). Geometry and structural plasticity of synaptic connectivity. *Neuron*, **34**(2), 275–288.
- Stephan, K. E., Baldeweg, T., and Friston, K. J. (2006). Synaptic plasticity and dysconnection in schizophrenia. *Biological psychiatry*, **59**(10), 929–939.
- Straatsma, T., Berendsen, H., and Stam, A. (1986). Estimation of statistical errors in molecular simulation calculations. *Molecular Physics*, **57**(1), 89–95.
- Stumpf, M. P., Wiuf, C., and May, R. M. (2005). Subnets of scale-free networks are not scale-free: sampling properties of networks. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(12), 4221–4224.
- Swanson, L. W. and Lichtman, J. W. (2016). From cajal to connectome and beyond. *Annual Review of Neuroscience*, **39**(1).

- Sylvester, J. J. (1878). Chemistry and algebra. *Nature*, **17**, 284.
- Tononi, G., Sporns, O., and Edelman, G. M. (1994). 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.
- Towlson, E. K., Vértes, P. E., Ahnert, S. E., Schafer, W. R., and Bullmore, E. T. (2013). The rich club of the c. elegans neuronal connectome. *The Journal of Neuroscience*, **33**(15), 6380–6387.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., and Joliot, M. (2002). Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *Neuroimage*, **15**(1), 273–289.
- Valles-Catala, T., Massucci, F. A., Guimera, R., and Sales-Pardo, M. (2016). Multilayer stochastic block models reveal the multilayer structure of complex networks. *Physical Review X*, **6**(1), 011036.
- van den Heuvel, M. P. and Fornito, A. (2014). Brain networks in schizophrenia. *Neuropsychology review*, **24**(1), 32–48.
- Van Den Heuvel, M. P. and Pol, H. E. H. (2010). Exploring the brain network: a review on resting-state fmri functional connectivity. *European Neuropsychopharmacology*, **20**(8), 519–534.
- Van Den Heuvel, M. P. and Sporns, O. (2011). Rich-club organization of the human connectome. *The Journal of neuroscience*, **31**(44), 15775–15786.
- van den Heuvel, M. P. and Sporns, O. (2013). Network hubs in the human brain. *Trends in cognitive sciences*, **17**(12), 683–696.
- van den Heuvel, M. P., Kahn, R. S., Goñi, J., and Sporns, O. (2012a). High-cost, high-capacity backbone for global brain communication. *Proceedings of the National Academy of Sciences*, **109**(28), 11372–11377.
- van den Heuvel, M. P., Kahn, R. S., Goñi, J., and Sporns, O. (2012b). High-cost, high-capacity backbone for global brain communication. *Proceedings of the National Academy of Sciences*, **109**(28), 11372–11377.
- van den Heuvel, M. P., Sporns, O., Collin, G., Scheewe, T., Mandl, R. C., Cahn, W., Goñi, J., Pol, H. E. H., and Kahn, R. S. (2013a). Abnormal rich club organization and functional brain dynamics in schizophrenia. *JAMA psychiatry*, **70**(8), 783–792.

- van den Heuvel, M. P., Sporns, O., Collin, G., Scheewe, T., Mandl, R. C. W., Cahn, W., Goñi, J., Hulshoff Pol, H. E., and Kahn, R. S. (2013b). Abnormal rich club organization and functional brain dynamics in schizophrenia. *JAMA psychiatry*, **70**(8), 783–92.
- van den Heuvel, M. P. V. D. and Sporns, O. (2011). Rich-Club Organization of the Human Connectome. **31**(44), 15775–15786.
- Van Dijk, K. R., Hedden, T., Venkataraman, A., Evans, K. C., Lazar, S. W., and Buckner, R. L. (2010). Intrinsic functional connectivity as a tool for human connectomics: theory, properties, and optimization. *Journal of neurophysiology*, **103**(1), 297–321.
- Van Dijk, K. R., Sabuncu, M. R., and Buckner, R. L. (2012). The influence of head motion on intrinsic functional connectivity mri. *Neuroimage*, **59**(1), 431–438.
- Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T., Bucholz, R., Chang, A., Chen, L., Corbetta, M., Curtiss, S. W., *et al.* (2012). The human connectome project: a data acquisition perspective. *Neuroimage*, **62**(4), 2222–2231.
- Varshney, L. R., Chen, B. L., Paniagua, E., Hall, D. H., Chklovskii, D. B., *et al.* (2011). Structural properties of the caenorhabditis elegans neuronal network. *PLoS Comput Biol*, **7**(2), e1001066.
- Vértes, P. E., Nicol, R. M., Chapman, S., Watkins, N., Robertson, D. A., and Bullmore, E. T. (2011). Topological isomorphisms of human brain and financial market networks. *Frontiers in systems neuroscience*, **5**, 75.
- Wang, J., Wang, Z., Aguirre, G. K., and Detre, J. A. (2005). To smooth or not to smooth? roc analysis of perfusion fmri data. *Magnetic resonance imaging*, **23**(1), 75–81.
- Wang, L., Metzak, P. D., Honer, W. G., and Woodward, T. S. (2010). Impaired efficiency of functional networks underlying episodic memory-for-context in schizophrenia. *The Journal of Neuroscience*, **30**(39), 13171–13179.
- Watts, D. J. (1999). *Small worlds: the dynamics of networks between order and randomness*. Princeton university press.
- Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of small-world networks. *nature*, **393**(6684), 440–442.
- Weisskoff, R., Baker, J., Belliveau, J., Davis, T., Kwong, K., Cohen, M., and Rosen, B. (1993). Power spectrum analysis of functionally-weighted mr data: whats in the noise. In *Proc. Soc. Magn. Reson. Med*, volume 1.

- Westbrook, C. and Roth, C. K. (2011). *MRI in Practice*. John Wiley & Sons.
- Woolrich, M. W., Ripley, B. D., Brady, M., and Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of fmri data. *Neuroimage*, **14**(6), 1370–1386.
- Xu, T., Yang, Z., Jiang, L., Xing, X.-X., and Zuo, X.-N. (2015). A connectome computation system for discovery science of brain. *Science Bulletin*, **60**(1), 86–95.
- Xu, X.-K., Zhang, J., and Small, M. (2010). Rich-club connectivity dominates assortativity and transitivity of complex networks. *Physical Review E*, **82**(4), 046117.
- Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., *et al.* (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of neurophysiology*, **106**(3), 1125–1165.
- Young, M. P. *et al.* (1992). Objective analysis of the topological organization of the primate cortical visual system. *Nature*, **358**(6382), 152–155.
- Yu, Q., Sui, J., Liu, J., Plis, S. M., Kiehl, K. A., Pearlson, G., and Calhoun, V. D. (2013a). Disrupted correlation between low frequency power and connectivity strength of resting state brain networks in schizophrenia. *Schizophrenia research*, **143**(1), 165–171.
- Yu, Q., Sui, J., Liu, J., Plis, S. M., Kiehl, K. a., Pearlson, G., and Calhoun, V. D. (2013b). Disrupted correlation between low frequency power and connectivity strength of resting state brain networks in schizophrenia. *Schizophrenia Research*, **143**(1), 165–171.
- Zalesky, A., Fornito, A., and Bullmore, E. T. (2010). Network-based statistic: identifying differences in brain networks. *Neuroimage*, **53**(4), 1197–1207.
- Zalesky, A., Fornito, A., Seal, M. L., Cocchi, L., Westin, C.-F., Bullmore, E. T., Egan, G. F., and Pantelis, C. (2011). Disrupted axonal fiber connectivity in schizophrenia. *Biological psychiatry*, **69**(1), 80–89.
- Zalesky, A., Fornito, A., and Bullmore, E. (2012a). On the use of correlation as a measure of network connectivity. *Neuroimage*, **60**(4), 2096–2106.
- Zalesky, A., Fornito, A., Egan, G. F., Pantelis, C., and Bullmore, E. T. (2012b). The relationship between regional and inter-regional functional connectivity deficits in schizophrenia. *Human brain mapping*, **33**(11), 2535–2549.
- Zalesky, A., Fornito, A., Cocchi, L., Gollo, L. L., van den Heuvel, M. P., and Breakspear, M. (2016). Connectome sensitivity or specificity: which is more important? *NeuroImage*.

- Zanghi, H., Ambroise, C., and Miele, V. (2008). Fast online graph clustering via erdős-rényi mixture. *Pattern Recognition*, **41**(12), 3592–3599.
- Zhang, Y., Lin, L., Lin, C.-P., Zhou, Y., Chou, K.-H., Lo, C.-Y., Su, T.-P., and Jiang, T. (2012). Abnormal topological organization of structural brain networks in schizophrenia. *Schizophrenia research*, **141**(2), 109–118.