THE UNIVERSITY *of* EDINBURGH

Novel Topological and Temporal Network Analyses for EEG Functional Connectivity with Applications to Alzheimer's Disease

Keith Malcolm Smith

Doctor of Philosophy

The University of Edinburgh

February 2018

THE UNIVERSITY *of* EDINBURGH
School of Engineering

Dedication

For my grandmother, Constance Clafferty.

Keith Malcolm Smith

February 2018

Abstract

This doctoral thesis outlines several methodological advances in network science aimed towards uncovering rapid, complex interdependencies of electromagnetic brain activity recorded from the Electroencephalogram (EEG). This entails both new analyses and modelling of EEG brain network topologies and a novel approach to analyse rapid dynamics of connectivity. Importantly, we implement these advances to provide novel insights into pathological brain function in Alzheimer's disease.

We introduce the concept of hierarchical complexity of network topology, providing both an index to measure it and a model to simulate it. We then show that the topology of functional connectivity estimated from EEG recordings is hierarchically complex, existing in a scale between random and star-like topologies, this is a paradigm shift from the established understanding that complexity arises between random and regular topologies. We go on to consider the density appropriate for binarisation of EEG functional connectivity, a methodological step recommended to produce compact and unbiased networks, in light of its new-found hierarchical complexity. Through simulations and real EEG data, we show the benefit of going beyond often recommended sparse representations to account for a broader range of hierarchy level interactions.

After this, we turn our attention to assessing dynamic changes in connectivity. By constructing a unified framework for multivariate signals and graphs, inspired by network science and graph signal processing, we introduce graph-variate signal analysis which allows us to capture rapid fluctuations in connectivity robust to spurious short-term correlations. We define this for three pertinent brain connectivity estimates- Pearson's correlation coefficient, coherence and phase-lag index- and show its benefit over standard dynamic connectivity measures in a range of simulations and real data.

Applying these novel methods to EEG datasets of the performance of visual short-term memory binding tasks by familial and sporadic Alzheimer's disease patients, we uncover disorganisation

of the topological hierarchy of EEG brain function and abnormalities of transient phase-based activity which paves the way for new interpretations of the disease's affect on brain function.

Hierarchical complexity and graph-variate dynamic connectivity are entirely new methods for analysing EEG brain networks. The former provides new interpretations of complexity in static connectivity patterns while the latter enables robust analysis of transient temporal connectivity patterns, both at the frontiers of analysis. Although designed with EEG functional connectivity in mind, we hope these techniques will be picked up in the broader field, having consequences for research into complex networks in general.

Lay summary

The brain is a vastly complex system in which great scientific efforts are being made to uncover its functional mechanisms. Its complex, dynamically changing nature requires high temporal resolution recording approaches such as the electroencephalogram (EEG). The EEG is also a particularly practical method for use in the detection and prevention of brain diseases and disorders at the earliest possible stages over the general population.

In this thesis, we will explore novel methods for modelling and capturing the architecture of brain function, as well as robustly capturing its dynamics, from direct recordings of the brain's electrical activity recorded at the scalp, i.e. non-invasive EEG. Since estimated brain function from EEG signals commonly takes the form of a network, it is naturally suited to analysis using techniques from the field of network science. To this end, I will introduce a number of novel network science methodologies designed with the EEG in mind.

It is shown that the hierarchical topology of EEG networks is strikingly complex compared to known network models. From this it is considered whether hierarchical complexity should play a role in determining the number of connections to analyse in the network. Whereas the literature generally recommends very few connections, the evidence suggests that a larger number of connections better captures the full range of complex network interactions.

We will then consider the problem of capturing dynamic connectivity. Since connectivity is usually defined as an average of a function over a period of time, being able to capture transient connectivity poses big problems in terms of finding a robust measure with high resolution. To address this, we will see that weighting instantaneous signal dynamics by long-term connectivity estimates provides a highly resolved, robust estimate of dynamic connectivity.

I will then apply these techniques of hierarchical complexity and dynamic connectivity to the problem of characterising Alzheimer's Disease (AD) at early stages and show that AD indeed has abnormal characteristics of hierarchical topology and dynamic connectivity, providing important insights into the pathology of this critical disease.

List of Symbols

- G is a graph.
- $\mathcal{V} = \{1, 2, \ldots, n\}$ is the node set of a graph.
- $\mathcal{E} = \{(i, j) : i, j \in \mathcal{V}\}$ is the edge set of a graph.
- $\mathbf{W}$ is a weighted graph adjacency matrix.
- w_{ij} is the entry of $\mathbf{W}$ corresponding to (i, j).
- $\mathbf{A}$ is a binary graph adjacency matrix.
- a_{ij} is the entry of $\mathbf{A}$ corresponding to (i, j).
- n is the size (number of nodes) of a graph.
- m is the number of edges of a graph.
- k_i is the degree of node i.
- F is a general function.
- $\mathbf{X}$ is a multivariate signal.
- $\mathbf{x}_i$ is a single time-series of $\mathbf{X}$.
- t is the index for time samples.
- T is the number of samples in a signal.
- $\bar{\mathbf{x}}_i$ is the mean value of $\mathbf{x}$.
- $P_{\mathbf{x}_i\mathbf{x}_j}$ is the cross-spectral density function.
- ω is a frequency.
- Ω is a frequency band.

- $s_i^a(t)$ is the instantaneous amplitude of $\mathbf{x}_i$.
- $\phi_i(t)$ is the instantaneous phase of $\mathbf{x}_i$.
- C is the global clustering coefficient of a graph.
- C_{loc} is the local clustering coefficient of a graph.
- L is the characteristic path length of a graph.
- E is network efficiency.
- Q is the modularity of a graph.
- q_i is the normalised degree of a graph.
- p_i is the proportion of nodes in a graph with the same degree as i.
- H is the entropy of a graph.
- $\mathbf{L}$ is the graph Laplacian.
- $\mathbf{D}$ is the degree diagonal matrix.
- s_i is the neighbourhood degree sequence of node i.
- $d_{i,j}$ is the jth element of s_i.
- D is the set of degrees of a graph.
- R is the hierarchical complexity of a graph.
- $\mathcal{D}$ is the number of distinct degrees of a graph.
- μ_{kj} is the mean value of element j over all k-length degree sequences of a graph.
- r_k is the number of nodes of degree k in a graph.
- $\mathbf{p}$ is the discrete cumulative distribution function parameter of the Weighted Complex Hierarchy (WCH) model.
- s is the additional weight parameter of the WCH model.
- l is the hierarchy level parameter of the WCH model.

- V is the degree variance of a graph.
- $\mathbf{k}$ is the degree sequence of a graph.
- P is the density of a graph.
- $\bar{G}$ is the weighted average of a set of graphs.
- $\mathbf{W}$ is the average of a set of weighted adjacency matrices.
- $K3$ and $K4$ are the complete graphs of size 3 and 4, respectively.
- $C_{\mathbf{w}}$ is the weighted clustering coefficient.
- LF is the leaf fraction of a graph.
- D is the diameter of a graph.
- MD is the maximum degree of a graph.
- Γ is a graph-variate signal.
- $F_{\mathcal{V}}$ is a general node function of a graph-variate signal.
- $F_{\mathcal{E}}$ is a general edge function of a graph-variate signal.
- $\bar{F}_{\mathcal{E}}$ is an edge dimension preserving function of a graph-variate signal.
- $\bar{F}_{\mathcal{V}}$ is a node dimension preserving function of a graph-variate signal.
- $\underline{\mathbf{J}}$ is a tensor of node functions for graph-variate signal analysis.
- $\underline{\boldsymbol{\Delta}}$ is a graph-variate network.
- $H_{\mathcal{V}}$ is function of connectivity of two signals.
- $\mathbf{C}$ is a graph connectivity adjacency matrix.
- θ is a Graph-Variate Dynamic (GVD) connectivity function.
- θ_i is a node GVD connectivity function.
- $\theta_{\mathcal{V}_a}$ is a module GVD connectivity function.

- $\theta_{\mathcal{V}_a,\mathcal{V}_b}$ is a between module GVD connectivity function.
- $z(t)$ is an autoregressive process.
- ϵ is white Gaussian noise.
- ρ is Pearson's correlation coefficient.
- $\hat{\mathbf{W}}$ is the weighted adjacency matrix with self-loops.
- τ is an epoch of a signal.
- μV is a micro volt.

List of Acronyms

- **AD** - Alzheimer's Disease
- **BOLD** - Blood Oxygen-Level-Dependent
- **CST** - Cluster-Span Threshold
- **CWN** - Complete Weighted Network
- **E-R** - Erdös-Rényi
- **ECO** - Efficiency Cost Optimisation
- **EEG** - Electroencephalogram
- **ERP** - Event-Related Potential
- **FDR** - False Discovery Rate
- **FIR** - Finite Impulse Response
- **fMRI** - functional Magnetic Resonance Imaging
- **GFT** - Graph Fourier Transform
- **GSP** - Graph Signal Processing
- **GVD** - Graph-Variate Dynamic
- **MCI** - Mild Cognitive Impairment
- **MDE** - Modular Dirichlet Energy
- **MEG** - Magnetoencephalogram
- **MMSE** - Mini-Mental State Examination
- **MST** - Minimum Spanning Tree

- **PLI** - Phase-Lag Index
- **SD** - Standard Deviation
- **USP** - Union of Shortest Paths
- **VSTM** - Visual Short-Term Memory
- **WCH** - Weighted Complex Hierarchy

Acknowledgements

This thesis would not have been possible without the guidance and supervision of Dr. Javier Escudero who gave ever generously of his time and care throughout my PhD studentship and who procured the funding to make it possible. Indeed, I would like to thank the EPSRC for providing that funding. I am grateful to Prof. John Starr and Prof. John Thompson for their wisdom and guidance as my second supervisors. I am grateful to Dr. Mario A. Parra for his willingness and energy throughout our collaboration and for providing the data of the visual short-term memory tasks used throughout this thesis. I would like to thank Dr. Daniel Abasolo for his feedback and help and for providing the EEG Alzheimer's disease dataset from Spain. I would also like to thank Dr. Chao Tan for agreeing to start a collaboration and his efforts to that end, and his student Ying Shen for her hard work. I would also like to thank all the editors and reviewers for their handling of manuscripts and their patience with my work and corrections. A special mention goes to my colleague Dr. Hamed Azami for our many discussions throughout the years and to Jonathan Mason, Dr. Mehrdad Yaghoobi, Dr. Athanasios Stavridis, Eli Kinney-Lang, Dr. Loukianos Spyrou, Dr. Steven Herbert, and many other members of the Institute past and present who joined us for fine dining at the Mayfield Bar and Grill and Swann Café.

I would like to thank the Royal Society of Edinburgh for funding my research visit to EPFL, Lausanne, Switzerland. I am grateful to Prof. Pierre Vandergheynst for hosting me during my visit to Lausanne, and especially to Dr. Benjamin Ricaud with whom I collaborated closely during my time there. A special mention goes to Francesco Grassi, Konstantinos Pitas, and Dr. Nauman Shahid for their friendship.

Finally, I am grateful to my partner Phung K.T. Nguyen, for her wisdom and support and lovingly sharing this experience with me, and to my family for their encouragement and support.

Declaration

I declare that this thesis has been composed solely by myself and that it has not been submitted, either in whole or in part, in any previous application for a degree. Except where otherwise acknowledged, the work presented is entirely my own.

Keith Smith

February 2017

Contents

Chapter 1

Introduction

1.1 Motivation

The work described here is motivated by the search to understand more about the dynamic interdependencies of brain function, particularly for practical applications in detecting pathology in the screening of the general population. The first step towards this was to explore state-of-the-art graph theoretic methods. In this exploration it became clear that there is much work still to be done in providing informative and robust tools to capture meaningful topological information. Thus, it is in graph topology where I take a large amount of focus in this thesis.

Amongst a wide range of problems from which this topic can benefit, one of the more pressing ones is the clinical problem of how to non-invasively and cost-effectively characterise and detect Alzheimer's Disease (AD) at an earlier stage than is currently possible. This is a problem of immediate and great importance due to the ageing population and the prevalence of AD in the elderly. To this end I apply my methods to Electroencephalogram (EEG) recordings of patients and healthy controls undergoing novel Visual Short-Term Memory (VSTM) tests. By targeting working memory functions involved in the early pathological decline of cognition, these tasks show promise in the sensitive and specific detection of AD. It is hoped that uncovering the functional underpinnings of task performance will go beyond providing insights into early pathology towards the uncovering of functional abnormalities before any outward symptoms can be seen. In the event of effective treatments of disease, success in this endeavour would provide enormous benefits to society in preventing brain damage associated with AD at earlier stages than is currently possible in the general population.

1.2 Objectives and Hypotheses

The main objectives of this thesis are:

- To aid in the construction of a concise and comprehensive network framework for EEG functional connectivity.
- Introduce powerful novel methods for topological analysis of EEG functional connectivity.
- Introduce powerful novel methods for temporal analysis of EEG functional connectivity.

It is hypothesised that by introducing novel complex network methods tailored to the EEG, we can better understand and extract information from the multivariate EEG activity.

Applying methods to EEG functional connectivity of AD, I expect to uncover novel information relating to abnormalities in AD functional connectivity.

1.3 Contributions

This thesis provides innovations in the topological and temporal analysis of networks.

For topology I introduce an entirely new concept of the hierarchical complexity of complex networks. I go on to provide an index to quantify hierarchical complexity and a generative model which traces a scale between networks with no topological hierarchy and networks with a strict topological hierarchy [1]. I then demonstrate the hierarchical complexity of EEG functional networks. I test the hypothesis that hierarchically complex weighted networks are best binarised in a medium density range, finding a non-arbitrary binarisation solution called the Cluster-Span Threshold which tends to dwell in this range, to be promising [2].

For temporal analysis, I introduce a novel approach, called graph-variate signal analysis, for assessing dynamic connectivity of networks constructed from multivariate signals. This begins with a unified framework for multivariate signals and network science from which graph-variate signal analysis is derived. I then formulate Graph-Variate Dynamic (GVD) connectivity for short-time dynamic connectivity estimation of multivariate signals [3] and provide formulae for the computation of modular GVD connectivity analysis [4]. These are

then applied in a number of simulated and real world experiments, including resting-state and task related EEG functional connectivity, finding GVD connectivity to be more robust and powerful than state-of-the-art approaches.

1.4 Structure of the Thesis

The remainder of the thesis is organized as follows:

Chapter 2: Review of the field. This chapter discusses the literature for brain networks, taking a focus on the networks of EEG functional connectivity. Various considerations which come up in the methodological pipeline are covered including connectivity estimation, binarisation, and state-of-the-art topological and temporal network analyses. Efforts made in AD research to date are also discussed.

Chapter 3: Hierarchical complexity of complex networks. The formulation, measurement and modelling of hierarchical complexity of binary networks is outlined. Its relevance to the network analysis of EEG functional connectivity, particularly in detailing the range of network density most appropriate for capturing hierarchical topologies, is then demonstrated.

Chapter 4: Accounting for hierarchical complexity in network binarisation. Following the new found knowledge of hierarchical complexity, I explore how this affects the way in which binarisation of the functional connectivity should be considered to provide the most meaningful and powerful binary network. It is shown that a hierarchically complex network is better characterised by larger densities than the sparse densities consistently recommended in the literature and that a complex hierarchical structure is more robust to targeted attacks of network hubs.

Chapter 5: Graph-variate signal analysis. The unified framework of multivariate signals and network science is formulated, from which graph-variate signals and their analysis are derived. I then define temporal connectivity for correlation, coherence and the phase-lag index and elaborate on a network science approach of graph-variate signals including node and modular components.

Chapter 6: Applications to visual short-term memory binding task performance and impairment

in AD. Hierarchy driven topological effects of task performance in Mild Cognitive Impairment (MCI) with high probability of progression to AD are determined. Dynamic functioning of healthy task performance in lateralised tasks are then located and it is found that these are of relevance to pathology in familial and sporadic AD in datasets involving binding tasks with a slightly different set-up.

Chapter 7: Discussion and conclusions. The main findings of the thesis are summarised followed by discussion of limitations and, finally, scope for future work is presented based on the novel developments herein.

Chapter 2

Review of the field

2.1 Introduction

In 1736 Leonard Euler, one of the most prolific mathematicians in history, provided a solution to a light-hearted puzzle: whether it was possible to traverse the seven bridges of the city of Königsberg exactly once each on a single walk [5]. The answer to this problem was in the negative. What is more important however, was that to solve it rigorously required him to invent a whole new branch of mathematics now known as graph theory. The term 'graph' here is not to be confused with the plotting of curves on an x-y axis. A graph is an object comprised of a set of points, which are commonly termed 'nodes', that share pairwise connections, typically drawn as lines in a dot-to-dot fashion, named edges [6]. By formulating the separate land masses of Königsberg as nodes and the bridges joining them as adjacent edges, Euler saw that the number of nodes with an odd number of adjacent edges needed to provide such a solution should be at most two since the path would need to leave each node, other than the starting and finishing ones, as many times as it arrived. This was not the case for Königsberg which had four landmasses all with an odd number of connecting bridges.

Since these humble beginnings, graph theory has evolved into a challenging branch of mathematics with famous problems such as the four colour theorem [7] and the graph isomorphism problem [8]. Even more recently, and arguably of more practical significance, graphs have been applied to deepen our understanding of the complex networks present throughout nature, providing new insights pertaining to structure and interdependency in problems of sociology [9], computer science [10], biology [11], economics [12], data science [13], and even linguistics [14], to name a few. In this thesis, we will focus on the

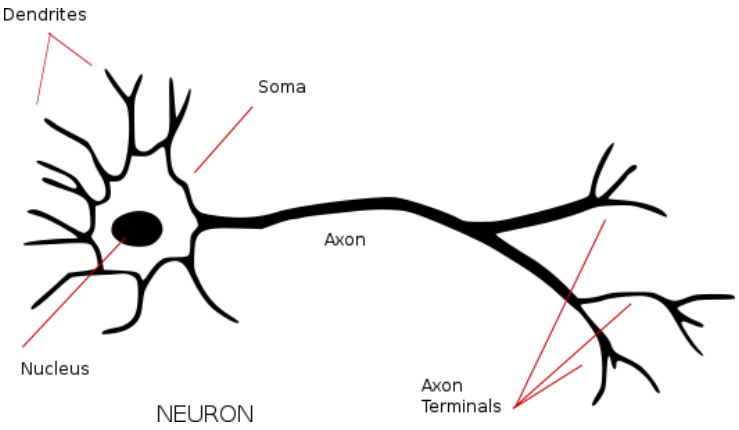

Figure 2.1: Illustration of a neuron with its various parts clearly labelled. Found online under a creative commons license [15].

development of methods for an exciting area in which the science of networks is helping to revolutionise the understanding of the subject– neuroimaging.

The brain is composed of billions of neurons (nerve cells). Each neuron has a central nucleus inside the cell body (soma), with a singular axon and numerous dendrites branching off like tentacles to make contact with other neurons, see Fig 2.1 for an illustration. This point of contact is called a synapse and interactions between neurons take place by means of electrochemical impulses. Each cell delivers this activity through its axon and receives it through its dendrites.

There are thus two major aspects to connectivity in neuroimaging– understanding the structure of physical connections between neurons, known as structural connectivity, and understanding the functionality of the brain through the inter-regional dependencies elicited through the electrochemical impulses, known as functional connectivity [16]. The function of the human brain– the way in which neurons, organised into function specific regions, interact with one another– is where we will focus our attention in this thesis. It is a vast topic of study whose importance is not only of an existential nature, in dealing with important questions such as how we think and perceive and how all the disparate functions of human consciousness are integrated to form a single seamless subjective experience, but its immediacy becomes apparent in the need to understand and tackle a wide array of pathological conditions including AD [17], epilepsy [18] and schizophrenia [19].

Recordings of whole-scale brain function generally come in the form of multivariate signals where each signal is associated to activity of localised, spatially separated neuronal clusters. These are recorded directly using sensors in the case of the EEG or the Magnetoencephalogram (MEG) or level of blood oxygenation, as the Blood Oxygen-Level-Dependent (BOLD) signal, in functional Magnetic Resonance Imaging (fMRI) [16].

2.2 The Electroencephalogram

The apparatus required to record the EEG consists of a cap holding a set of electrodes placed at usually evenly spaced distances over the scalp, as shown in Fig 2.2, according to a standardised system, such as the 10-20 system [20]. It directly records the aggregated voltages of electrical activity from the brain, eliciting the use of reference electrodes due to the relativity of voltage. EEG recordings provide a unique opportunity to deepen our understanding of human brain function across a healthy lifespan and in diseases of the nervous system [21]. In the clinical context, the low cost, practicality and portability of the EEG offers a strong feasibility for screening purposes in contrast to the MEG and fMRI which require expensive hi-tech stationary equipment and special magnetically shielded and radio frequency shielded rooms, respectively. The EEG can thus aid in the early detection of brain dysfunction associated to diseases which have an impact on the worldwide population, such as dementia [22–24]. In a broader context, the high temporal resolution of the EEG presents a great opportunity to study the rapid interdependent processes which underlie cognition [21]. Thus, the EEG provides an unparalleled matching of practicality and data richness for neurological diagnostics.

That being said, it would be amiss not to recognise certain limitations of the EEG. Particularly, it has low spatial resolution, which is compounded by being unable to reliably solve the inverse problem to retrace the activity at the scalp to accurate brain sources. Apart from the fact that the number of sources at any moment in time is unknown, which is necessary to accurately solve such a problem, volume conduction through the irregular biological tissues between brain and scalp surface causes unpredictable and non-negligible errors in localisation [25]. However, this by no means deters us from extracting meaningful features of brain activity from EEG and particularly in understanding complex dynamic relationships of brain function which can be captured using network science methods at the sensor level [16, 26, 27]. In this

Figure 2.2: Picture of an EEG cap worn by a man seated in front of a computer. Courtesy of Dr. Mario A. Parra, Psychology Department, Heriot-Watt University.

case we are not so much interested in the specifics of where activity is occurring, which, for example, might typically be the case if we wished to identify a region of abnormal excitatory activity giving rise to seizures suitable for ablation, but rather in characterising the topology of statistical dependencies between more general brain regions. This does mean, however, that results obtained need to be analysed with care.

2.3 Estimates of EEG functional connectivity

Network topologies of function from EEG recordings are established by assigning each signal to a node and implementing a bivariate analysis on all signal pairs which are encoded as weights of the relevant adjacent edges, see Fig 2.3 for an illustration. This is often preceded by filtering the EEG signal in a frequency band of interest. A very common suggestion, based on spectral analysis, are frequency bands of Delta (0-4Hz), Theta (4-8Hz), Alpha (8-13Hz), Beta (13-32Hz), and Gamma (32-60Hz), although the exact interval of the bands can vary.

Formally, we define $G = (\mathcal{V}, \mathcal{E}, \mathbf{W})$ as a graph where $\mathcal{V} = \{1, 2, \ldots, n\}$ is the node set, $\mathcal{E} = \{(i, j) : i, j \in \mathcal{V}\}$ the edge set and $\mathbf{W}$ the weighted adjacency matrix with entries w_{ij} the weight (estimated connectivity strength) of edge (i, j) for $i \neq j$ and $w_{ii} = 0 \, \forall i \in \mathcal{V}$. The edge density of a network, $P = 2m/n(n-1)$, normalises the number of edges in a network with respect to its size. For additional details of the important basic concepts of network science see Fig 2.4.

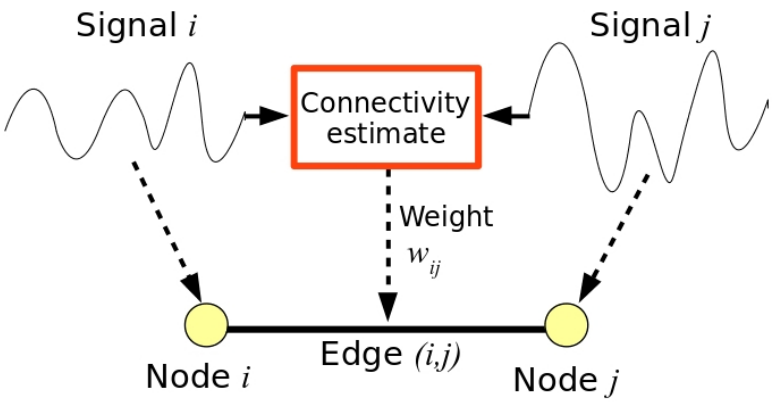

Figure 2.3: Illustration of the construction of a graph from multivariate brain signals. Signal i and j are mapped to nodes i and j of a graph. The output of connectivity estimation between these signals becomes the weight of the edge between them.

The weighted functional connectivity adjacency matrix, $\mathbf{W}$, is generally computed as the output of a function, F, on the multivariate signal $\mathbf{X} \in \mathbb{R}^{n \times T}$ as

$$\mathbf{W} = F(\mathbf{X}) \tag{2.1}$$

and

$$w_{ij} = F(\mathbf{x}_i, \mathbf{x}_j), \tag{2.2}$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ are the individual time-series, i.e. rows, of $\mathbf{X}$ related to nodes i and j of the graph. Specifying the function F is a challenging research area with many proposed solutions. Taking $\bar{\mathbf{x}}_i$ as the mean of $\mathbf{x}_i$ over epoch T, a basic attempt would be to take F as Pearson's correlation coefficient:

$$F(\mathbf{x}_i, \mathbf{x}_j) = \frac{\sum_{t=1}^{T}(x_i(t) - \bar{\mathbf{x}}_i)(x_j(t) - \bar{\mathbf{x}}_j)}{\sqrt{\sum_{t=1}^{T}(x_i(t) - \bar{\mathbf{x}}_i)^2}\sqrt{\sum_{t=1}^{T}(x_j(t) - \bar{\mathbf{x}}_j)^2}}. \tag{2.3}$$

This tracks the behavioural similarities of the signal amplitudes, where large positive values are attained when both signals instantaneously go in the same direction, relative to their means, and large negative values attained going in opposite directions. However, signal amplitudes of nearby electrodes are highly susceptible to spuriously high values due to volume conduction [25], illustrated in Fig 2.5.

Coherence is another proposed measure which works with the power and cross-power spectral

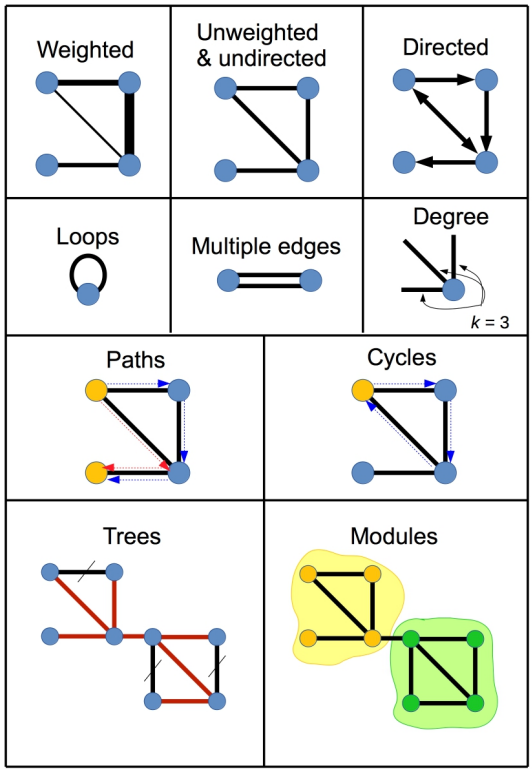

Figure 2.4: Illustration of basic network science concepts. Networks can be weighted (top left) where wider edges indicate larger weights, or directed (top right) where arrows indicate direction of edges, but we are interested mostly in unweighted and undirected graphs (top centre). Networks can also have loops (2nd row, left) and multiple edges (2nd row, middle), but these are generally not allowed in functional connectivity networks. The degree of a node (2nd row, right) is the number of edges adjacent to it. A path (3rd row, left) is a sequence of unique edges from which one can get from a given node to another. A cycle (3rd row, right) is a path which starts and ends at the same node and a tree (bottom left) is a graph with no cycles. Modules (bottom right) are subsets of nodes which are often chosen in a way that the within module nodes are highly connected and less connected to the rest of the graph.

density functions, $P_{\mathbf{x}_i\mathbf{x}_j}(\omega)$, for $i = j$ and $i \neq j$, respectively, and frequency ω [28]. For a frequency band of interest, Ω, we write

$$F(\mathbf{x}_i, \mathbf{x}_j, \Omega) = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} \frac{|P_{\mathbf{x}_i\mathbf{x}_j}(\omega)|^2}{P_{\mathbf{x}_i\mathbf{x}_i}(\omega) P_{\mathbf{x}_j\mathbf{x}_j}(\omega)}. \tag{2.4}$$

This assesses correlations of the frequency components of time-series, but is similarly effected by volume conduction as correlation. Indeed, since volume conduction through the head is practically speaking instantaneous to all electrodes it will always be present as synchronised

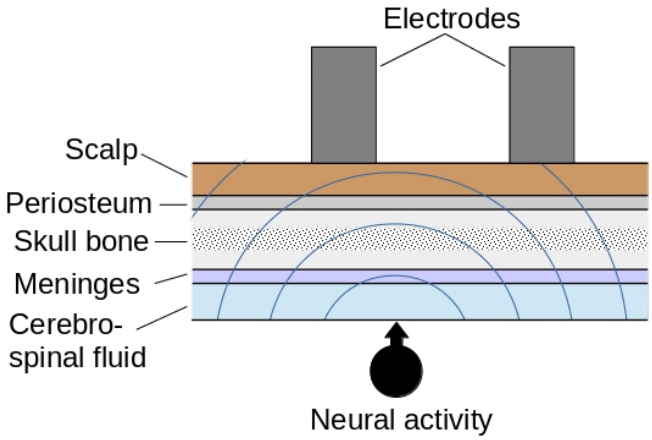

Figure 2.5: Illustration of volume conduction of neural electrochemical activity through head tissue and fluid. Each medium has its own conductivity. In this example, the neural activity is closer to the left electrode but will be picked up strongly in the right electrode also.

activity and thus any computation based on amplitudes or frequency synchronisations will be corrupted by volume conduction.

Instead, more promising approaches look at non-synchronised frequency dependencies. These are much more likely to be true dependencies since they cannot be directly accounted for by volume conduction. There are a number of different formulations of such measures, for a comprehensive list see [29] or [30] for example. Here, we will focus on the Phase-Lag Index (PLI) [31] which has been consistently successful in providing insights into AD, particularly, from EEG/MEG signals [32–35].

Recall that the analytic signal of $\mathbf{x}_i$ is composed of instantaneous amplitude and phase components $x_i^a(t) = s_i^a(t)e^{j\phi_i}(t)$, where $s_i^a(t)$ is the envelope of the signal t and $\phi_i(t)$ is the instantaneous phase at time t. The PLI is defined as

$$F(\mathbf{x}_i, \mathbf{x}_j) = |\langle \text{sgn}(\phi_i(t) - \phi_j(t)) \rangle|, \tag{2.5}$$

i.e. the magnitude of the average over time of the signed values of differences of the instantaneous phases of signals $\mathbf{x}_i$ and $\mathbf{x}_j$. This measures the consistent phase differences between time-series, indicating lead/lag dependencies. Generally, this is taken as an undirected value since determining whether one signal is truly leading or lagging another is an issue which requires careful consideration. For example, one could study the expected

delays between inter-regional signal transmissions to infer plausible phase dependencies. Estimation here could be based on Euclidean distances between sensor placements or on fibre tract trajectory of structural connectivity data. One would also need to consider indirect dependencies from longer transmission pathways. Such involved processes are beyond the scope of this thesis and we will be content just now with looking at undirected PLI.

2.4 Binarising functional connectivity

Whichever connectivity measure is chosen, functional connectivity defined between all possible pairs of signals presents the researcher with a full adjacency matrix whose entries are only distinguished by relativity of weights. The most popular network science techniques are based on binary networks [36], where in the notation above $\mathbf{W}$ is replaced with binary adjacency matrix $\mathbf{A}$ with entries $a_{ij} \in \{0, 1\}$ indicating non-existence (0) or existence (1) of an edge. This led initial efforts on the analysis of brain network topologies to be implemented via the binarisation of the weights using some arbitrarily chosen thresholds with some good success [16, 37, 38].

Still, studying the original weighted networks holds appeal in that it is more direct and has advantages of maintaining the information of relative strengths of connections [39]. But serious complications exist in that these computed weights are known to vary due to any number of different pre-processing choices or connectivity analyses implemented, thus complicating comparisons and obfuscating results [40, 41]. Furthermore, since the weights of dependency measures generally follow a non-scaling distribution between 0 and 1, many interesting topological considerations in binary networks, such as concerning paths, become redundant in light of the fact that the shortest weighted path between any two nodes is likely to be just the weighted edge connecting them [42]. Therefore, binarising the weights remains a more widely used approach which can explain the main topology of the underlying activity while alleviating methodological biasing and topological redundancy of weights.

Arbitrarily defining edges is both wide open to bias, where researchers can pick and choose the best threshold for their particular dataset, see Fig 2.6, and makes comparisons between studies intractable since many network analyses are dependent on the network's number of edges. Taking the most reasonable route that the strongest connections are those which should

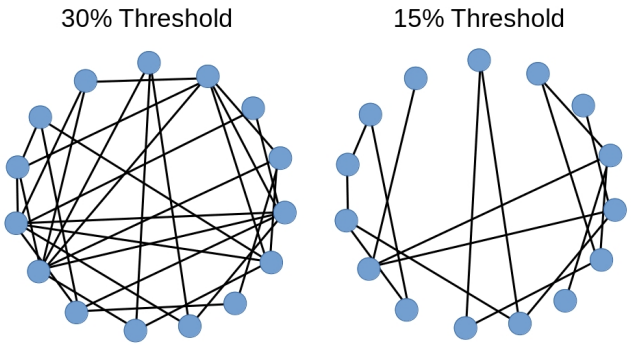

Figure 2.6: Illustration of how the chosen threshold affects the network topology and thus introduces subjective bias to the research. Left is an example of a network with 30% of strongest connections kept and right with 15% of strongest connections kept.

be kept and weakest discarded, the problem is best illustrated by realising that the total possible number of thresholds to choose from is equal to the number of edges, which, in a complete, weighted, undirected network, is $n(n-1)/2$.

Network binarisation is thus an important active problem in brain network research. Two main approaches exist to non-arbitrarily define the threshold. These are statistical methods and topological methods. In the former case, the problem is posed as finding those edges which are statistically likely to be true connections. However, rather than resolving arbitrary choices, we note that this merely diverts it towards the statistical significance paradigm, where arbitrary standards (e.g. significance level $\alpha = 0.05$ in widespread statistical tests) have long been adopted to mitigate an intractable problem. Further problems with a statistical approach relate to difficulties in finding the correct solutions for the numerous available connectivity measures in a way that is consistent and reliable, biases from the size of available data, and, in the case of data surrogate methods, biases due to network size [26]. Topological approaches to binarisation, on the other hand, have attracted attention recently, providing criterion for binarisation which are not subject to arbitrary threshold decisions and inherently complement the analysis of network topology.

One method gaining traction in the field is the Minimum Spanning Tree (MST) [43,44]. A tree in graph theory is a graph with no cycles, see Fig 2.4. A spanning tree of a graph is a connected tree subgraph that includes all nodes of the graph. The MST is thus the spanning tree of a weighted graph that requires the least total weight. Fortunately, the algorithm to construct the MST is quite simple and is included in popular toolboxes, e.g. in MATLAB [45]. For

functional connectivity, we are not interested in the minimum weight, but in fact the maximum weight (strongest connectivity), though the nomenclature of using 'minimum' has stuck [43]. The algorithm is exactly the same regardless– one adds the strongest connectivity weights to the network one by one; if at any point a cycle is created, the edge is discarded and we move on to the next one; the algorithm ends when a spanning tree has been constructed [46].

Another method is the Union of Shortest Path graphs (USP) [47]. The shortest path between two nodes in a network is the set of edges with the minimum sum of weights connecting them. This can be constructed using Dijkstra's [48] algorithm to find the shortest paths between each pair of nodes in the network, adding all the edges of those paths to an initially empty binary network. Because connectivity has an inverse relation to distance, the weights of the network must first be relationally inverted in order to construct the shortest paths. This inversion process can take several forms which involves a certain amount of subjective discretion and depends largely on the distribution of the original weights.

The Cluster-Span Threshold (CST) chooses the binary network at the point where open to closed triples are balanced [49]. A triple is a path of length two, containing three nodes and two edges. A closed triple, or triangle, is a path of length three which begins and ends on the same node, containing three nodes and three edges, and an open triple is a triple which is not closed. A closed triple, $\{(i,j),(j,k),(k,i)\}$, thus contains three triples– $\{(i,j),(j,k)\}$, $\{(j,k),(k,i)\}$ and $\{(i,k),(k,j)\}$– and a complete graph is the only graph in which all triples achieve closure. The number of closed triples is calculated from the sum of the diagonal of $\mathbf{A}^3$ (number of paths of length 3 beginning and ending at the same node), and the number of triples is calculated by summing all non-diagonal elements of $\mathbf{A}^2$ (diagonal elements being the number of 'paths' of length two consisting of traversing the same edge twice– $\{(i,j),(j,i)\}$ and non-diagonal elements, a_{ij}, being the number of triples starting at i and ending at j). The balance of open and closed triples is thus assessed by the fraction of triples which are closed. Importantly, the balancing of this topological characteristic necessarily endows the binary network with a trade-off of sparsity to density of edges. We see this since a network is sparse if most triples are open and dense if most triples are closed. It is hypothesised that this balance achieves an informational richness useful for capturing different topologies of EEG functional connectivity [42]. For the case of fMRI connectivity, where structural connectivity is

known, it has yet to be considered whether matching the clustering coefficient to the structural data could achieve a more appropriate threshold than the CST.

The final threshold we shall look at is the Efficiency Cost Optimisation (ECO). The ECO proposes a threshold to keep the strongest $1.5n$ edges– equivalent to a density threshold of $3/(n-1)$– which is an approximation based on consistent observations of simulated and real brain networks of the maximum ratio of the combined local and global network efficiencies and density [50]. It is hypothesised that such a trade-off of network efficiency and sparsity provides networks which are meaningful to the concept of economy in brain function [51].

2.5 Network topology of functional connectivity

Once the network has been binarised, the network is ripe for a wide array of topological analyses both at the level of individual nodes (local indices) and whole-scale properties of the network (global indices). Initial studies of brain network topology derive from the seminal work by Watts and Strogatz [52] which described small-world characteristics of a network as having a high clustering coefficient, interpreted as segregation of a network into clusters, and a low characteristic path length, interpreted as a well integrated network.

Two forms of the clustering coefficient exist, the global clustering coefficient, C, and the local clustering coefficient, C_{loc}. C is the ratio of closed to open triples in the network, depicted in Fig 2.7, which can be formulated, for a network with at least one triple, as

$$C = \frac{\sum_{i,j,k=1}^{n} a_{ij}a_{jk}a_{ki}}{\sum_{i,j,k=1}^{n} a_{ij}a_{jk}} \tag{2.6}$$

where a_{ij} are the entries of the binary adjacency matrix $\mathbf{A}$. C_{loc} is, instead, the average over nodes of the triples centred at a given node which achieve closure and reflects local properties of clustering. This can be written as

$$C_{loc} = \frac{1}{n} \sum_{i \text{ s.t. } k_i > 1} \frac{\sum_{j,k=1}^{n} a_{ij}a_{jk}a_{ki}}{\sum_{j,k=1}^{n} a_{ij}a_{ik}}, \tag{2.7}$$

where k_i is the number of edges adjacent to i, called the degree of i. The fraction within the first summand is the local clustering coefficient of node i, C_{loc}^{i}.

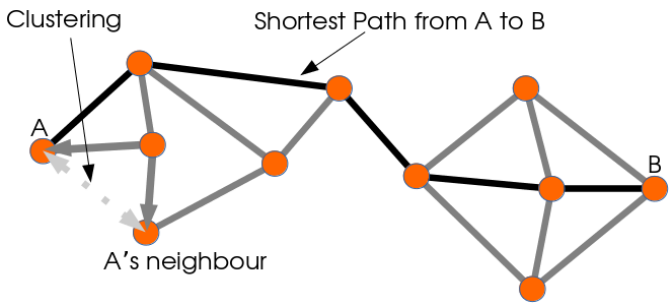

Figure 2.7: Illustrations of clustering and shortest paths in a network. If node A shares an edge with A's neighbour, the formerly open triple, whose edges are denoted by the dark grey arrows, becomes closed. The shortest path from A to B is the path with the smallest number of edges (not necessarily unique) starting at A and finishing at B, indicated by the black edges.

The characteristic path length, L, is the average of the shortest paths, depicted in Fig 2.7, between all pairs of nodes. It can be written

$$L = \frac{1}{n(n-1)} \sum_{i,j=1}^{n} d(i,j), \tag{2.8}$$

where $d(i,j)$ is the number of edges in the shortest path between nodes i and j. This is related also to network efficiency defined as

$$E = \frac{1}{n(n-1)} \sum_{i \neq j=1}^{n} \frac{1}{d(i,j)}, \tag{2.9}$$

i.e. the average inverse shortest path length, which is set to measure the efficiency of connectivity pathways in the network.

Adopted in the very first studies of functional connectivity networks [53, 54], these measures have found small-world properties suggesting a trade-off of integrative and segregative behaviour in brain function [16, 55] allowing for 'economic' operations [51].

Another seminal paper in complex networks by Barabási and Albert [56] noted that the degree distributions of real-world networks tend to be close to scale-free and following up on this it was found that this property also appears in functional brain networks [57, 58]. Essentially, the scale-free property of the degree distribution means that there are a few very high degree nodes whereas most nodes have small degrees. In fact, for sparse networks found in many real-world situations, there can only be a few very high degree nodes, as the number of edges

is limited, thus perhaps the more interesting aspect of a scale-free distribution is the existence of these high degree nodes. In brain networks, these nodes are called hub nodes [59] and are seen as key components allowing for the integration and distribution of information flow, since the brain is essentially a large information flow control system [60].

These preliminary studies provided important insights into the topology of functional brain networks, however, in our own analysis, we found that they can be refined and in Chapter 3 we will discuss how the concepts of integration/segregation and hub dominance can be captured using single indices, paying particular attention to EEG functional connectivity.

Another key aspect of network topology is that of community structure, introduced by Newman and Girvan [36, 61]. The community structure of a network refers to the decomposition of the network into highly intra-connected modules of nodes which are relatively weakly interconnected. Modularity is defined as the difference between the number of within module edges in the network with those predicted by random network with the same degree distribution, written as

$$Q = \frac{1}{2m} \sum_{i,j=1}^{n} \left(a_{ij} - \frac{k_i k_j}{2m} \right) \delta(\mathcal{V}_i, \mathcal{V}_j), \tag{2.10}$$

where $\mathcal{V}_i$ is the module containing node i and $\delta()$ is the Kronecker delta function which is 1 if $\mathcal{V}_i = \mathcal{V}_j$ and 0 otherwise. This then requires predefined modules. The general strategy for an arbitrary network is to find the maximal value of Q over all possible non-trivial module decompositions of the network. Efficient algorithms have been created [61, 62] aiming to do this, attaining reasonable results. Applied to functional connectivity, modular organisation has indeed been found in functional connectivity networks [63, 64], suggesting brain function is facilitated through the integration of specialised functional modules.

These three aspects of network topology– integration/segregation, scale-freeness and modularity– have been proposed as the main 'dimensions' of graph topology [65] and due to successful findings in brain networks, this has been brought forward for functional connectivity [55, 66].

2.6 Complexity of network topology

The inherent complexity of networks, rather than being an independent aspect in and of itself, is presumed to arise as a result of networks finding a trade-off between orderly structures and the efficiency provided by random topologies [51], see Fig 2.8.A. We do not take this view. We propose that the complexity of a network is more intrinsically dependent on the hierarchical structure of node degrees, related to the relationship of network hubs and lower degree nodes. Particularly, neither random nor regular networks have any hub nodes, whereas hubs are one of the main characteristics of real-world networks. Chapter 3 details methods which substantiate this proposition. We devise a novel network index called hierarchical complexity to assess the complexity of a network based on hierarchical levels. We then create a model whose main parameter, s, goes from random networks at $s = 0$ to strict class-based networks at $s = 1$, see Fig 2.8.B. Applying this to EEG resting-state data, we seek to assess the hierarchical complexity of EEG functional connectivity compared to the model and whether or not complexity can be seen to arise from either the random to regular spectrum hypothesis or the hierarchical spectrum hypothesis of equal node networks to class-based networks. Note that, with respect to brain networks, the concept of hierarchy can be applied in several ways, e.g. [64, 67], but in this work we consider the hierarchy of network degrees and the inter-level relationships of this hierarchy.

We note that network entropy is an index which has been used to assess complexity of the network [65]. This is defined using the normalised degree distribution $q_i = k_i p_i / \sum_j k_j p_j$, where k_i is the degree of node i and p_i is the proportion of nodes in the graph with the same degree as node i which relates to probabilities of 'going to'/ 'coming from' neighbouring nodes in directed graph problems. Then the entropy of graph G is a straightforward derivation of Shannon's entropy equation [68] for the degrees of the graph:

$$H = -\sum_{i=1}^{n} q_i \log(q_i), \tag{2.11}$$

encoding the eccentricity of the graph degrees. We compare hierarchical complexity with this measure and show its greater performance in identifying different network types.

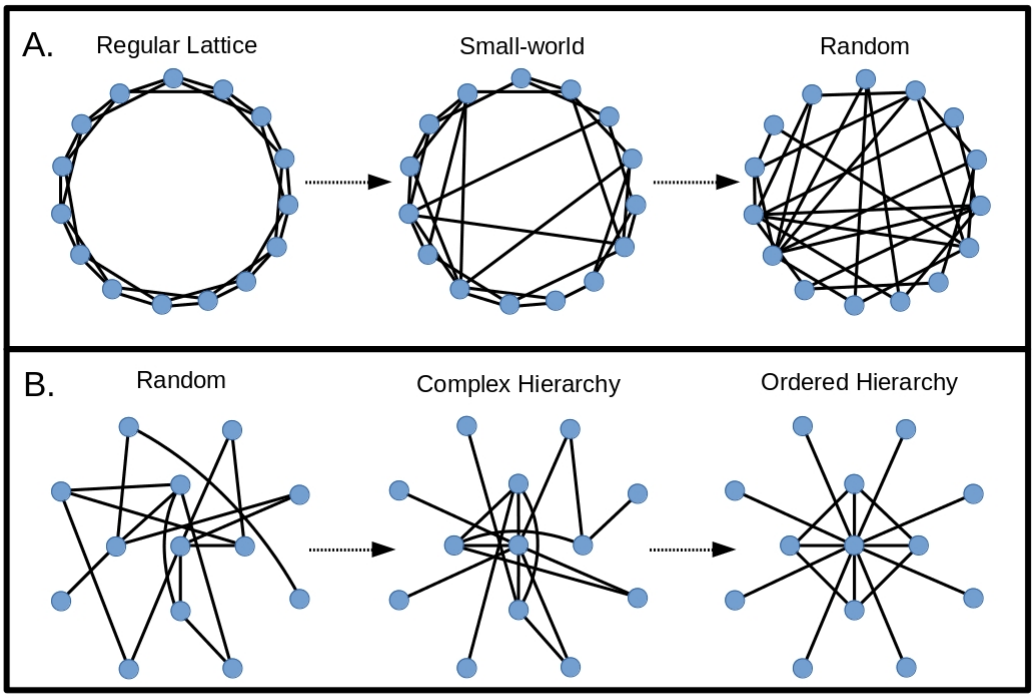

Figure 2.8: A. The previous model of network complexity: randomly rewiring a regular lattice one obtains a small-world network before a complete rewiring returns a random graph. B. The proposed model of network complexity: starting from a random graph, one adds additional weight, s, to the adjacent edges of chosen nodes. For $s < 0$ a complex hierarchical topology emerges, whereas $s \geq 1$ returns an ordered hierarchy. Notice that the complex hierarchical model emerges from randomness, whereas the small-world model emerges from orderliness.

The hierarchical complexity of network topology brings forth important considerations of how the weighted functional connectivity network should be binarised. Sparse representations such as the MST and ECO seek a minimalistic model based on the sparsity found in structural connectivity [38]. However, considering a complex hierarchical structure, such sparsity would conceivably severely limit topological information by being dominated by interactions of hub nodes and disregarding more nuanced connectivity information available throughout the hierarchy. We hypothesise that for more subtle changes in functional connectivity, such as relating to different cognitive tasks or early stages of diseases, sensitivity will be enhanced by considering denser networks that contain more information of a network's degree hierarchies. We explore this hypothesis in Chapter 4 and also look at the effect of network attacks– compromises to the network weights which aim to simulate noise and/or functional

degeneration– on the sensitivity of different binarisation approaches in picking up differences in populations of differently configured network models.

2.7 Dynamic connectivity

Static network topologies of functional connectivity allow us to assess local and whole-scale stable interdependencies of brain regions. However, one of the major aspects of brain function is its transient, dynamic nature [69]. For instance, in a single cognitive task lasting only a few seconds, several functionally important and distinct periods may be recognised which last only a fraction of a second each, e.g. pre-stimulus, stimulus, retention, test, response (e.g. [70]). It is expected that each of these periods will require or induce separate functional activities and inter-regional relationships. Therefore, the importance of developing dynamic network models of functional connectivity, particularly for high temporal resolution EEG, is paramount for understanding the complexities of brain function [26, 27, 69].

A large contingent of research solutions for temporal networks take the form of events occurring at edges (i.e. between two nodes) which change over time, geared towards data in which node specific activity is either not available or not meaningful [71]. Such outputs are also well suited to a multi-layer network framework where chronologically separated networks can be integrated as layers into a multi-layer network with a tensor adjacency matrix in which topological considerations, such as closure of triples, can span the layers [72].

Initial attempts have done well to devise temporal and multi-layer network methods to analyse multivariate signals where most recent studies go the route of implementing disjoint [73] or overlapping [74–76] windows to construct a number of distinct chronologically separated graphs to gain a foothold on changing connectivity patterns. However, this is limited by the length of the window– the less samples used to define the network, the less reliable is the connectivity estimate. Fig 2.9 (a) illustrates this, showing independent realisations of an autoregressive process in which spurious strong correlations can be found in short windows. On the other hand, the larger the window used the less meaningful it is at determining temporally refined connectivity estimation. Therefore obtaining reliable transient information is difficult.

Another study proposes the signals' instantaneous phase differences as a measure of

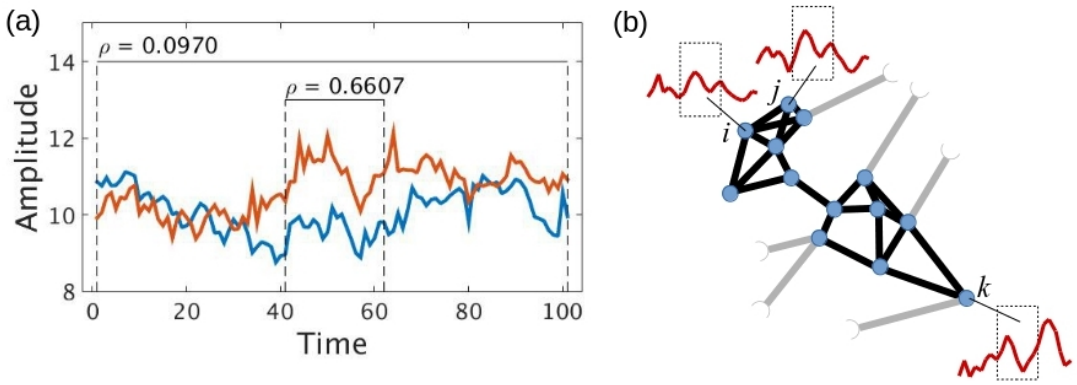

Figure 2.9: (a) Specific example of spurious short-term correlation coefficient, ρ, from independent realisations of an autoregressive model. (b) Illustration of how long-term connectivity weighting (black edges connecting blue nodes) improves robustness of analysis of short-term transient dynamics. Nodes i, j and k all exhibit similar behaviour in the windowed epoch. However, from the topology of long-term connectivity it is clear that the correlation between i and j, with a shortest path of 1, is more meaningful whereas correlations between i/j and k, with shortest paths of 6, are spurious and should be disregarded.

instantaneous connectivity [77]. Though this allows signal resolution analysis, it is still wide open to spurious connections, Fig 2.9 (a), and noisy fluctuations.

Another promising framework for exploring the possibility of dynamic connectivity analysis is found in Graph Signal Processing (GSP). In GSP, a signal, whose samples occur at graph nodes, is processed over the graph topology. However, GSP is mainly concerned with the development of a cohesive signal processing theory for graph signals, analogous to classical signal processing [78]. Spectral graph techniques are implemented, using the eigen-decomposition of either the graph adjacency matrix [79] or its Laplacian [78], to process graph signals in a method called the Graph Fourier Transform (GFT) which has been applied in topics such as big data [80] and neuroscience [81]. Recent work on the integration of the temporal domain within the GSP framework is also under way [82, 83]. This spectral approach, however, presents hurdles in interpretation in light of the fact that the frequencies of the graph signal emerge through graph eigenvectors which relate to a still unquantified extent to the graph topology. Further, the graph signal itself remains a passive component in the analysis treated as a vector separate from the graph adjacency matrix.

On the other hand, the Dirichlet energy of a graph signal, another component of the GSP framework, is a more directly extracted feature which weights instantaneous activity by graph

weights [78]. This is defined as

$$\mathbf{x}^T\mathbf{L}\mathbf{x} = \sum_{i,j=1}^{n} w_{ij}(x_i - x_j)^2 \tag{2.12}$$

for graph weights w_{ij} and graph signal $\mathbf{x}$, where $\mathbf{L}$ is the graph Laplacian defined as $\mathbf{D} - \mathbf{W}$ where $\mathbf{D}$ is the matrix whose diagonal is the degrees of $\mathbf{W}$ and zero elsewhere. In chapter 5 we will describe the advancement of this perspective of Dirichlet energy towards a general dynamic connectivity estimate for functional connectivity based on a novel *graph-variate signal analysis* framework which unifies multivariate signals and graphs in a fully flexible way. The advantage of this method for EEG functional connectivity is that, by weighting the short-term connectivity estimates by stable long-term dependencies, a more robust estimate can be realised. The illustration in Fig 2.9.(b) provides a useful intuitive grasp of this, where instantaneous connectivity between nodes i and j is of high interest whereas that between i/j and k is likely to be spurious based on the information obtain from stable connectivity estimates indicated by the edges.

2.8 Application to Alzheimer's Disease

To show the power and new insights offered by the novel methods described in this thesis, in Chapter 6 we apply them to data retrieved for research into AD. AD is a neurodegenerative disease whose physical hallmarks are the presence of toxic amyloid plaques [84] and neurofibrillary tangles of Tau proteins [85] originating in memory related regions such as the entorhinal cortex and the hippocampus but eventually progressing throughout the whole brain in a process lasting many years. These toxic plaques cause neuronal deaths which instigate initial problems of memory loss, leading to the complete decimation of brain function and eventual death.

Two important forms of AD are the sporadic version, occurring in elderly individuals and set to be a major problem in the future due to the increasingly elderly population, and the familial version which occurs with certainty in specific inherited genes with onset in middle age [86]. Of course, many genetic variants have been identified that increase the risk of sporadic AD, but unlike those for familial AD, these in themselves are not sufficient to cause the disease.

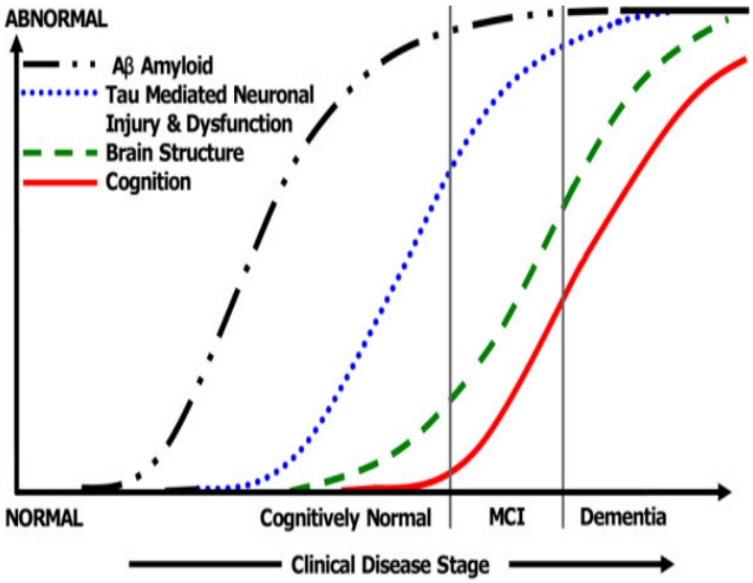

Figure 2.10: Diagram of the progression of AD and of current markers for its detection. This model derives from familial AD, but is thought to apply to sporadic AD also. Clearly, early, pre-clinical markers are necessary to help in the hopeful prevention of irreversible whole-scale damage to the brain. Found in [87].

Since the aetiological pathway is clear in familial AD and it is also possible to predict who will develop the disease many years before symptoms of the disease occur, familial AD is often adopted as a model for sporadic AD. It is commonly thought that people with sporadic AD usually pass through a phase where memory deficits are not too severe and where the impact of these on daily social and occupational functioning is minimal which is labelled as MCI. However, not all people who fulfil diagnostic criteria for MCI (i.e. have mild cognitive deficits) progress to AD. Since treatments are thought to be most effective before symptoms of AD have appeared, identifying those with MCI who will progress to AD is of great current importance. Especially considering that by this point the physical damage done by the disease is already substantial, Fig 2.10.

Because of the associated neuronal deaths it is proposed that early, non-invasive signs of the disease may be found in disruptions to the brain's functional architecture via synaptic failure [88]. Following this, studies of functional connectivity have been conducted to validate this hypothesis. Initial studies of resting-state EEG functional connectivity of AD focused on small-world characteristics, finding deviations from healthy ageing [32,37,89] that were related to a loss of complexity and efficiency. Progressively, investigations have looked closer into the

key role of the deterioration of network hubs in corrupting functional integration in pathology [32, 90–92]. Other important associated factors of hierarchical network topology such as loss of assortativity [89]– neighbouring degree correlations– and loss of hub connectivity to distant nodes [93, 94] have been found. Since these aspects relate directly to the hierarchy of network degrees, where hubs are top level nodes and assortativity describes the strength of bonds within hierarchy levels, it is of interest to look more closely into the details of network topology vis-à-vis hierarchical structure. Furthermore, using the EEG as a tool for screening purposes over the general population is very promising since it is non-invasive, cost-efficient and practical.

Honing in on initial stages of the disease related to memory loss or even before MCI, a promising explorative approach based on EEG functional connectivity relates to performance of short-term memory cognitive tasks. Evidence suggests that the binding of shape and colour in VSTM [95] is depreciated in both sporadic AD [70, 96] and familial AD [97] compared to performance of VSTM of single features (shape or colour alone). Following this, we will explore topological and temporal analysis of EEG functional connectivity during VSTM binding in familial MCI developing AD with certainty and elderly MCI at high-risk of AD. In doing so we seek to assess the physiological underpinnings of the expected dysfunction in AD causing poor task performance and thus take steps towards possible early EEG biomarkers of this disease.

This will provide important evidence of the utility of the introduced methods to a major societal problem and we believe this can be extended to various clinical diseases such as schizophrenia, epilepsy and Parkinson's as well as more broad understanding of functional connectivity during various cognitive processes and conscious or unconscious states.

Chapter 3

The Complex Hierarchical Topology of EEG Functional Connectivity

The contributions of this chapter were published in the Journal of Neuroscience Methods in January, 2017 [1].

3.1 Introduction

Complexity is understood neither to mean regularity, where obvious patterns and repetition are evident, nor randomness, where no pattern or repetition can be established, but attributed to systems in which patterns are irregular and unpredictable such as in many real world phenomena [98]. Particularly, the brain is noted to be such a complex system [99]. Here, hierarchical complexity is concerned with understanding how the degree hierarchy of the network contributes to its complexity.

We introduce a new index aptly named hierarchical complexity, R, which is based on targeting the structural consistency at each hierarchical level of network topology. Alongside this, we introduce the Weighted Complex Hierarchy (WCH) model which simulates hierarchical structures in weighted networks. This model works by modifying uniform random weights by addition of multiples of a constant, which is essentially a weighted preferential selection method with a highly unpredictable component provided by the original random weights. We show that it follows very similar topological characteristics of networks formed from EEG phase-lag connectivity. Intrinsic to our model is a strict control of weight ranges for hierarchical levels which offers unprecedented ease, flexibility and rigour for topological comparisons in applied settings and for simulations in technical exploration for brain network analysis. This

also provides an alternative to methods which randomise edges [52, 53] or weights [39] of the original network.

Any rigorous evaluation of brain networks should address their inherent complete weighted formulation [26]. However, the current field has largely lacked any concerted effort to build an analytical framework specifically targeted at Complete Weighted Networks (CWNs), preferring instead to manipulate the functional connectivity CWNs into sparse binary form (e.g. [44, 53, 100] as well as wide-spread use of the Watts-Strogatz [52] and Albert-Barabasi [56] models) and using the pre-existing framework built around other research areas– such as social science and the internet– which have different aims and strategies in mind [6]. In our methodological approach we propose novel generalisations of pre-existing sparse binary models to CWN form and thus allow a full density range comparison of our techniques. Due to the intrinsic properties of these graph types we find minimal and maximal topologies which can help to shed light on a wide variety of topological forms and their possible limitations [65] in a dense weighted framework.

Further, as part of our study we seek after straightforward indices to evaluate other main aspects of network topology for comparisons [65, 66] and, in this search, found it necessary to revise key network concepts of integration-segregation [41, 45, 52] and scale-freeness [57, 101]. We provide here these revisions: i) that the clustering coefficient, C, is enough to analyse the scale of integration and segregation, finding it unnecessary and convoluted to use the characteristic path length, L, as a measure of its opposite as previously proposed [16, 52]; ii) we provide mathematical justification that the degree variance, V, and thus network irregularity [102] is a strong indicator of the scale-free factor of a topology.

3.2 Network science: proposed methods and key revisions

3.2.1 Hierarchical Complexity Index

The ideas of order and complexity are well known in the discussion of networks (indeed, real world networks are often called complex networks [16, 27, 103]). In mathematics, the graphs studied derive from some theoretical principles. These can involve set patterns, without random fluctuations of edges, such as regular networks, fractal networks, star networks and

grid networks. On the other hand much interest is shown in more randomly generated topologies, such as random graphs and other graphs involving random processes, as these express something of the more erratic and irregular quality of edges in networks constructed from real world phenomena [52,104]. However, such phenomena differ from random processes in that there is a clear organisational behaviour apparent throughout the structure [64] and there is a distinct presence of very highly connected nodes [56], presenting a hierarchy of node degrees. Although this structure is perhaps impossible to retrace, because its formation inevitably involves many unknown generative processes, we can provide methods for its analysis.

Studying the relationships of node degrees has provoked interesting findings in brain networks. For example it has been found that a small group of highly connected nodes create a rich club [103] and nodes with generally lesser connectivity exist on a peripheral lower levels. Further, it is seen that a node's relationship within the context of the network is greatly determined by the other nodes to which it is connected [105]. Thus, to understand the hierarchical complexity of a network we propose to study the behaviour of nodes of a given degree by looking at the degrees of nodes in their neighbourhoods. Instead of determining centralities of nodes as in eigenvector centrality, for example, we are interested in the variability of connections found in nodes with the same level of centrality in order to understand the complexity of organisational principles in the network. We define D as the set of degrees of a graph, G. Similar to the idea of node degree sequences [106], we can construct neighbourhood degree sequences specific to each node in the graph. That is, for a node i of degree $k \in D$ we have a sequence

$$s_i = \{d_{i,1}, d_{i,2}, \ldots, d_{i,k}\} \text{ s.t. } d_{i,1} \leq d_{i,2} \leq \cdots \leq d_{i,k} \in D,$$

where $d_{i,j}$ is the degree of the jth node connected to node i (see Fig 3.1.A). For all nodes of a given degree, k, the corresponding neighbourhood degree sequences have equal length, k.

We define the hierarchical complexity, R, of a network as the average variance of the k-degree neighbourhood degree sequences and can be expressed as:

$$R = \frac{1}{\mathcal{D}} \sum_{r_k > 1} \frac{1}{k(r_k - 1)} \left(\sum_{j=1}^{k} \left(\sum_{i \in D_k} (s_{ki}(j) - \mu_{kj})^2 \right) \right), \tag{3.1}$$

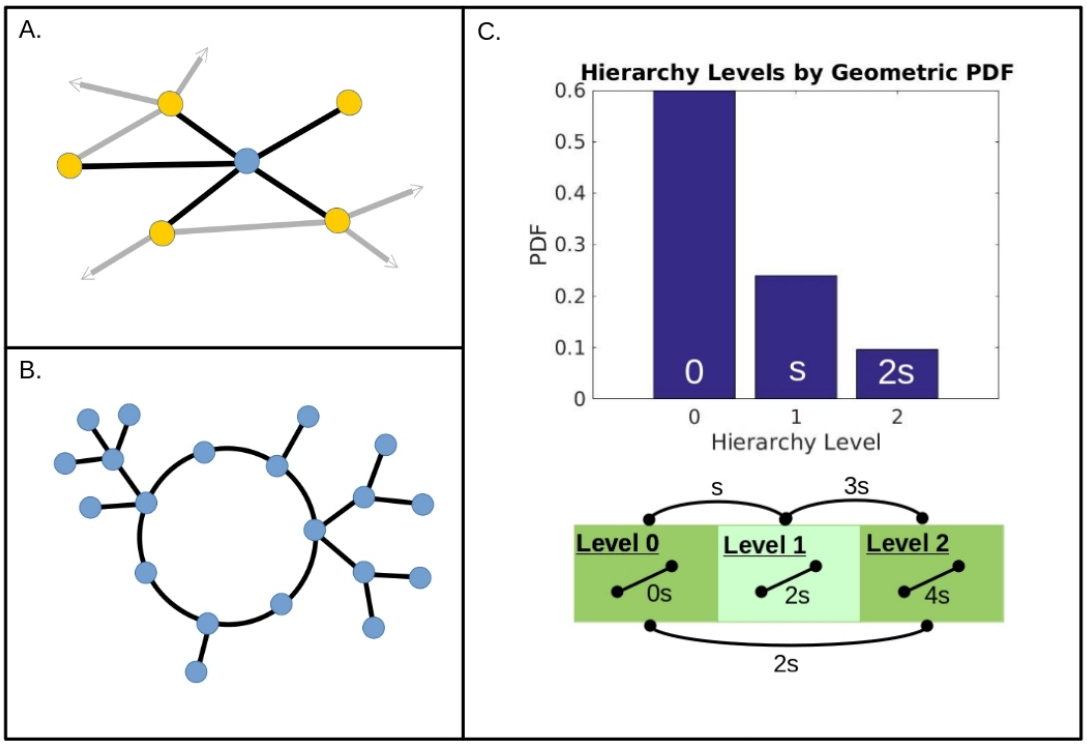

Figure 3.1: A. Example of a node degree neighbourhood. Here is shown a part of a network relating to the neighbourhood of the blue node. The blue node has neighbourhood degree sequence $\{1, 2, 3, 4, 4\}$*, i.e. the ordered degrees of the orange nodes. Grey edges indicate all the additional edges of the orange nodes in the network. B. Example for graph complexity. Here is shown a 20 node network with varying ordered-ness at different degree levels. C. Diagram of the construction of the WCH model. Above is the probability distribution function for a geometric distribution with* $p = 0.6$ *for a three level hierarchy. Below is a graphic displaying the additional weight added between nodes in given hierarchy levels.*

where $\mathcal{D}$ is the number of distinct degrees in the graph, D_k is the set of nodes of degree k, $s_{ki}(j)$ is the jth element of the ith k-length sequence, μ_{kj} is the mean value of element j over all k-length sequences and r_k is the number of nodes of degree k.

Organisation of the graph at the level of k-degree nodes can be seen by comparing the jth elements of their neighbourhood sequences. If all of the jth elements of all the sequences are equal, that is $s_i = s_j$ for all s_i, s_j of length k, then there is a high degree of order present in the k-degree nodes of the graph. If these sequences differ widely however, then it can be said that the k-degree nodes are either disorganised or more complexly organised. For example, in Fig 3.1.B the two degree nodes all have the same degree sequences– $\{3, 4\}$– whereas the three degree nodes are split into two different degree sequences– $\{1, 2, 2\}$ and $\{1, 1, 4\}$– and finally

the neighbourhood degree sequences of the four degree nodes are all different– $\{1,1,1,4\}$, $\{1,2,2,4\}$ and $\{2,3,3,3\}$. So the complexity of just the two degree nodes is 0, the complexity of just the three degree nodes is $2((2-1.5)^2+(1-1.5)^2+(2-3)^2+(4-3)^2)/(4\times3\times3) = 5/36$ and the complexity of just the four degree nodes is $(2(1-4/3)^2 + (2-4/3)^2 + 2((1-2)^2 + (3-2)^2) + 2(4-11/3)^2 + (3-11/3)^2)/(4\times3\times2) = (16/3)/24 = 8/36$, the complexity over all three levels being the average: 13/108.

This measure is thus minimal (0) for graphs in which, for each k and k', every k-degree node is connected to exactly the same number of k'-degree nodes. This property, for example, is seen in ring lattices, and quasi-star graphs and is close to minimal in the line graph, fractal graphs and grid lattices. Furthermore, the degrees of random networks are known to have a fairly small spread which is a factor penalised by our complexity value. Thus random networks should obtain low values of our complexity measure. On the other hand, R values of real networks are expected to be higher given the high spread and degree fluctuations of those networks caused by hub nodes promoting a high degree irregularity while the spontaneous nature of real-world edges should promote a high variability of the neighbourhood degree sequences. This measure can be used in networks with leaf nodes, isolated nodes and disconnected components without error and maintaining the same interpretation. Note that this index is as yet not fully normalised and thus it can be expected that one may attain larger values of R in larger networks and it is certainly possible to achieve values of R greater than 1. It is hoped that normalisation is possible and future work will be undertaken to determine a suitable formula. Until then, caution is advised when comparing R between networks of different sizes.

3.2.2 Weighted Complex Hierarchy Model

The foundation of our model is the random CWN model. The most general random network is the Erdös-Rényi (E-R) random network [104] which is formed by assigning a probability, p, to the question of the existence or non-existence of edges on a network with n nodes. Such a construct is, in fact, an ensemble of graphs denoted $G(n,p)$. A sample of this ensemble is obtained by generating a random value for every possible edge and applying the probability value p as a threshold to see whether or not that edge should exist in our sample. The random CWN model is thus simply a symmetric matrix with zero diagonal and randomly generated values $w_{ij} \in [0,1]$ elsewhere. If we threshold the CWN at weight p, we recover a binary

Erdös-Rényi random graph from the random graph ensemble $G(n, p)$.

Starting from an Erdös-Rényi CWN we randomly distribute the nodes into hierarchy levels based on some discrete cumulative distribution function, $\mathbf{p}$, by generating a random number, r, between 0 and 1 for each node and putting the node in the level for which $r - \mathbf{p}$ is first less than 0. We then distribute $(l-1)s$ additional weight to all edges of adjacent nodes in the lth level, for some suitably chosen s. The parameters of this model are then $(n, s, l, \mathbf{p})$. The parameter n is the number of nodes in the network. The parameter s is the strength parameter, which is constant since the random generation of the initial weights is enough to contribute to weight randomness. The parameter l is the number of levels of the hierarchy, with a default setting of a random integer between 2 and 5. The vector $\mathbf{p}$ is the cumulative probability distribution vector denoting the probabilities that a given node will belong to a given level where the default, which we use here, is a geometric distribution with $p = 0.6$ in hierarchical levels $(1, 2, \ldots, l)$ where the nodes with highest connectivity (top hierarchical level) are at the tail end of the distribution. Fig 3.1.C plots an example of the geometric distribution for a three level hierarchy. The text inside the box plots, above, indicates the additional weights given to edges adjacent to nodes inside the given level. The graphic below that explains the additional weights provided by the strength parameter of edges between nodes in different levels as well as in the same level. For example, an edge between a level 1 node and a level 2 node has additional strength $3s$ which consists of one s provided by the node in Level 1 and $2s$ provided by the node in Level 2. At $s = 0$, we have the E-R random network and at $s = 1$ the weights of the network are linearly separable by the hierarchical structure producing a strict 'class-based' topology. Between these values a spontaneous 'class-influenced' topology emerges.

3.2.3 Revision of concepts from network science

Here we present justifications for indices as measures of key topological factors– the global clustering coefficient, C, for degree of segregation and the degree variance, V, for irregularity, linked to scale-freeness.

3.2.3.1 Integration-Segregation

The concept of integration in brain networks is closely tied in to the small world phenomenon [107], where real world networks are found to have an efficient 'trade off' between segregative and integrative behaviours [51]. The most widely used topological indices in network science– C and the characteristic path length, L– are commonly noted as measures of these quantities, respectively.

Since integration implies a non-discriminative behaviour in choice, we argue that the random graph ensemble [104], defined by its equal probability of existent edges between all pairs of nodes, is the most exemplary model of an integrated network. Anything which deviates from equal probability is a discriminative factor which favours certain edges or nodes over others, likely leading to more segregated activity. Further, it is clear that integration and segregation are opposite ends of the same spectrum– something which is not integrated must be segregated and vice versa. Having one index to inform on where a network lies on that spectrum is therefore sufficient.

Thus, here we propose C as the topological measure to evaluate levels of integration (and so segregation) of a given network. Firstly, we note that values of C for random graphs and small-world graphs are often much more distinguishable than those of L [52] and it is certainly assumed that these graphs have very different levels of integration. Secondly, since the random network is optimally integrated and $E[C_{\text{ran}}] = E[P_{\text{ran}}]$ [6], where P_{ran} is the edge density of the random network, then the larger the deviation from 1 of the value $\gamma = C/E[C_{\text{ran}}] = C/E[P_{\text{ran}}] = C/P$, the more segregated is the network. We will include both L and C in our analysis in order to provide evidence to back the above proposal.

3.2.3.2 Regularity and Scale-Freeness

Another topological factor of small world networks is noted as a scale-free nature characterised by a power law degree distribution [101]. To understand this aspect of network topology another factor of network behaviour is formulated distinguishing between line-like and star-like graphs [17, 41].

Here, we show that characterisation of scale-freeness is closely connected to the regularity of a

network. Regular graphs have been studied for over a century [108]. They are defined as graphs for which every node has the same degree. An almost regular graph is a graph for which the highest and lowest degree differs by only 1. Thus a highly irregular graph can be thought of as any graph whose vertices have a high variability. Such behaviour can be captured simply by the variance of the degrees present in the graph, that is

$$V = \text{var}(D), \tag{3.2}$$

where $D = \{k_i\}_{i \in \mathcal{V}}$, is the set of node degrees on a given graph [102].

For regular graphs $V = 0$ by definition, but more probing is necessary to distinguish high V topology. For a graph with degrees $\mathbf{k} = \{k_1, k_2, \ldots, k_n\}$, and $\sum_{i=1}^{n} k_i = 2m$, on multiplying out the brackets V simplifies to

$$\begin{aligned} V &= \frac{1}{(n-1)} \sum_{i=1}^{n} \left(\frac{2m}{n} - k_i \right)^2 \\ &= \frac{\|\mathbf{k}\|_2^2}{(n-1)} - 2mP, \end{aligned}$$

where $P = 2m/n(n-1)$ is the edge density and $\|\mathbf{k}\|_2^2 = \sum_{i=1}^{n} k_i^2$, is the squared ℓ_2 norm of $\mathbf{k}$. This tells us that V is proportional to the sum of the squares of the degrees of the graph, $\|\mathbf{k}\|_2^2$, and, for fixed number of edges, m, V in fact depends only on $\|\mathbf{k}\|_2^2$. Now, it is known that $\|\mathbf{k}\|_2^2$ is maximal in quasi-star graphs and quasi-complete graphs [109]. Essentially, the quasi-star graph has a maximal number of maximum degree nodes in the graph for the given edge density and the quasi-complete graph has a maximal number of isolated, or zero-degree, nodes in the graph. This tells us that, for low P, high V denotes the presence of a few high degree nodes and a majority of relatively low degree nodes, i.e. scale-free-like graphs. Thus, due to the restriction placed on possible degree distributions by the number of edges (the small number of edges in sparse networks means the number of high degree nodes is very limited), the irregularity of degrees is a strong indicator of the strength of decay of the given distribution, relating to how scale-free the graph is.

3.2.4 Complete Weighted Network Archetypes

Here we detail the method to generalise sparse binary network archetypes to CWN form. Suppose we have a set of unweighted graphs $G_1 = (\mathcal{V}, \mathcal{E}_1, \mathbf{A}_1)$, $G_2 = (\mathcal{V}, \mathcal{E}_2, \mathbf{A}_2)$, $\ldots$, $G_q = (\mathcal{V}, \mathcal{E}_q, \mathbf{A}_q)$, where $\mathcal{V}$ is a common vertex set, $\{\mathcal{E}_i\}_{i=1}^q$ are the edge sets and $\{\mathbf{A}_i\}_{i=1}^q$ are the adjacency matrices, such that $G_1 \subset G_2 \subset \cdots \subset G_q$. Then we define the weighted average form of the set of graphs $\{G_i\}_{i=1}^q$ as

$$\bar{G} = (\mathcal{V}, \mathcal{E}_q, \bar{\mathbf{W}})$$

where $\bar{\mathbf{W}}$ is a weighted adjacency matrix defined as

$$\bar{\mathbf{W}} = \frac{1}{q} \sum_{i=1}^{q} \mathbf{A}_i.$$

Particularly, if G_q is the complete graph then $\bar{G}$ is a complete weighted graph. In fact we can always arbitrarily choose G_q as such since all binary graphs on the same vertex set are contained within the complete graph on that vertex set. We call $H_i = \mathbf{A}_i - \mathbf{A}_{i-1}$ the ith weight category of $\bar{G}$, i.e. the adjacency matrix corresponding to all the edges existing in G_i which are not present in G_{i-1}. We say that H_i is the stronger weight category to H_{i-1}, reflecting the fact that those edges correspond to the ones with larger weights in $\bar{G}$.

We now present the complete weighted graphs for the regular ring lattice, star, grid lattice and fractal modular networks (see Fig 3.2 A,B,C,D respectively).

3.2.4.1 Regular Ring Lattice

A regular ring lattice is a network which we can illustrate by evenly spacing nodes in a circle and connecting each node to its k closest neighbours, giving a regular graph of degree k (Fig 3.2.A). Note that k must be an even number since equal spacing on a circle means that closest nodes come in pairs. The exception to this is when $n - 1$ is odd and $k = n - 1$ forms the complete graph. The regular ring lattice is then defined by the parameters (n, k). Some special examples are the closed triple with $(3, 2)$ and the regular ring lattice with parameters $(n, 4)$, which was presented by Watts and Strogatz to represent regular networks for comparison with

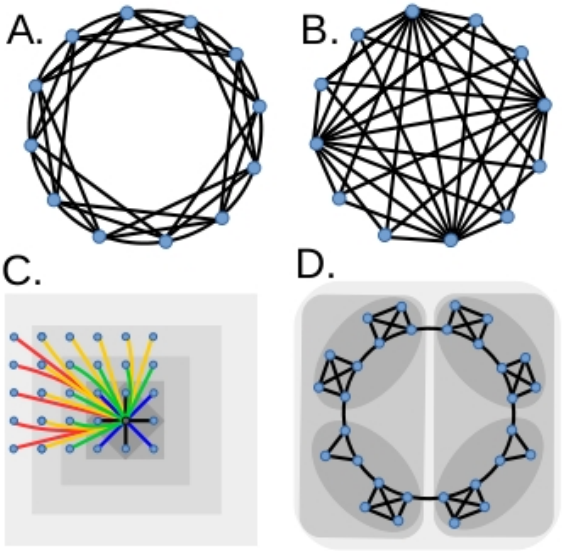

Figure 3.2: A. A 12 node ring lattice of degree 6, comprising the three strongest weight categories of the ring lattice CWN. B. The quasi-star with 4 nodes of degree $n-1$ and $n-4$ nodes of degree 4, also comprising the first four categories of the star CWN. C. The grid lattice weight categorisation (relating to the grey node) in a 30 node network (see supplementary material). Colours of edges denote category: black, blue, green, orange and red edges are in weight categories 1, 2, 3, 4 & 5, respectively. The increasingly lighter boundaries thus represent 'catchment' areas around the node by increasing category. Centring these 'catchment' areas around a given node gives the respective categorisation of edges adjacent to the new node. D. Fractal modular CWN weight categorisation on 30 nodes. Edges shown (black) are 1st weight category edges. In this instance, increasingly lighter background represents areas within which all pairs of nodes become connected by edges when the network is subject to the threshold corresponding to the respectively increasing category (see supplementary material in [1]).

small world networks [52].

The CWN for the ring lattice is thus formed with binary graphs $\{G_i\}_{i=1}^{n/2}$ defined as the regular ring lattice with parameters $(n, 2i)$ for $i = 1, \ldots, n/2$ when $n-1$ is even. When $n-1$ is odd the binary graphs are $\{G_i\}_{i=1}^{(n-1)/2}$, the complete network K_n being added to the end of the sequence as the largest graph.

3.2.4.2 Star

A star graph can be thought of as the archetypal scale-free-like graph with one node sharing edges to every other node and no other edges. Thus it has one node of degree $n-1$ and $n-1$ nodes of degree 1. We can construct a complete weighted generalisation of the star graph by

taking the classic star as G_1 and the subsequent graphs $G_2, G_3, \ldots, G_{n-1}$ associated with the increasingly higher density quasi-star graphs, i.e. G_i being the quasi star graph with i nodes adjacent to every other node in the network. G_{n-1} is the complete graph, where the $(n-1)$th node connects to the nth node to complete it. Eventually we have a CWN with $n-1$ weight categories consisting of $n-1, n-2, \ldots, 2, 1$ edges, respectively. Fig 3.2.B, shows the edges corresponding to the first four weight categories in a star CWN.

3.2.4.3 Grid Lattice

Another common lattice graph is the grid lattice where nodes are placed at the intersection of lines in a grid. This graph is a strictly 'class-based' graph by which we mean nodes can be separated into clearly distinct classes based on degree– corner nodes ($k = 2$), side nodes ($k = 3$) and central nodes ($k = 4$). We use this as an archetype for a distance based graph where closer nodes are more strongly connected. In order to construct a complete weighted graph following from the grid lattice topology we propose to categorise the edges as shown in Fig 3.2.C. This graph is similar to the regular lattice in that the nodes have stronger edges to nodes they are closer to. However weight categorisation by closeness is instead best represented by placing square 'catchment' areas around the nodes (Fig 3.2.C). Each category then consists of the edges within the corresponding catchment area placed around every node minus all the edges in the previous category. This results in an inhomogeneous number of edges at each node in a given category, creating a hierarchy of nodes based on degree, contrasting with the regular ring lattice where all nodes have equal degree by definition.

3.2.4.4 Fractal Modular

In order to obtain an ordered graph with a highly modular topology, we define here methods for constructing fractal modular graphs from some number of $K3$s and $K4$s, the complete graphs on 3 and 4 nodes, respectively. Here, we simply connect these $K3$ and $K4$ subgraphs in a ring as shown in Fig 3.2.D. All integers above 5 can be expressed as a sum of 3s and 4s so this method can be used to construct a graph with any $n > 5$. These networks are fractal because at each step the smaller modules merge into larger modules until we eventually have a complete graph, these steps are shown in Fig 3.2.D by the increasingly lighter grey backgrounds where nodes within the shaded area indicate that edges exist at that category

level. To select a $3,4$-summation of n as well as the ordering of module forming at each step we can simply use our discretion for graphs with a fairly low number of nodes. Here, for 30 nodes we choose six $K4$s and two $K3$s connected in a ring to construct the first weight category and progressive module forming as depicted in Fig 3.2.D. For 64 nodes we choose an initial weight category consisting of sixteen $K4$s connected in a ring with similar progressive module forming. Generally, the higher the power of 2 which is a factor of the initial combined number of $K3$ and $K4$ modules, the better the $3,4$-summation it is for the fractal composition of the network.

3.3 Methods

Here we apply methods to graphs of 64 nodes, typical of medium density EEG, to validate the developments in Section 3.2. For analysis we employ edge density thresholds at integer percentages of strongest weighted edges, rounded to the nearest whole number of edges. We then compute indices for each of these binary networks and plot the obtained values on a curve against edge density, similar as in e.g. [37, 110]. This generates index curves plotted against edge density which provides a detailed analysis of the CWN topology. Other methods exist to analyse CWNs such as weighted indices [39] or density integrated indices [111], but these indices still give only singular values for a given network which belies little of the topological behaviour at different scales of connectivity strength.

For random and WCH CWNs we use sample sizes of 100 for each network and for the EEG functional connectivity CWNs we have a sample size of 109 [112]. On the index curves for these we plot the median with the interquartile range shaded in. For ordered networks there is only one network per type by definition.

Our analytical framework is composed of a mixture of entirely new concepts and novel generalisations of existing concepts to CWN form. It is constituted of the following elements: four indices, R, C, V, Q characterising four important and distinct topological features; five CWN archetypal models– Random, Star, Regular Lattice, Fractal Modular, Grid Lattice, and the WCH model.

3.3.1 Indices

In [65, 66] an 'architecture' of network topology is proposed involving the three most widely studied properties of brain networks– integration (and segregation) [41, 45, 52], scale-freeness [57, 101] and modularity [36, 64]. For our analysis in comparison with hierarchical complexity, R, we choose a straightforward index for each of these topological factors– C for integration, V for scale-freeness and Q for modularity [36].

3.3.2 Comparisons for the WCH model

We implement comparisons with the Watts-Strogatz small-world model [52] which randomly rewires a set proportion of edges starting from a regular lattice. We use the full range of parameters for initial degree specification (2 up to 62) and random rewiring parameters from 0.05 in steps of 0.05 up to 0.95. For each combination of parameters, 100 realisations of the model were computed and C, V, Q, and R were measured.

We also compare with Albert-Barabasi's scale-free model [56] which begins with a graph consisting of a core of highly connected nodes to which the rest of the nodes are added one by one with a set degree but paired by edges to randomly selected nodes. We use an initial number of 15 nodes and the additional node's degree from 3 up to 14 in order to reach larger densities, since the size of the core limits the number of edges a new node can be adjacent to and thus limits the possible density.

3.3.3 EEG networks

We use an eyes open, resting EEG data set with 64 nodes. We report on networks created from the Beta band (13-32Hz) using coherence and the PLI in order to account for different possible types of EEG networks while reducing redundancy of similar topological forms found between the frequency bands (see supplementary material of [1]).

The dataset, recorded using the BCI2000 instrumentation system [113], was freely acquired from Physionet [112]. The signals were recorded from 64 electrodes placed in accordance with the international 10-10 system and are sampled at 160Hz. We took the eyes open resting state condition data, consisting of 1 minute of continuously streamed data which were partitioned

into 1s epochs and averaged for each of 109 volunteers.

FieldTrip [114] was used for pre-processing where the 64 channels were re-referenced using an average reference. The signals were then filtered with an order 40 Finite Impulse Response (FIR) filter with hamming windows in Beta before computing the connectivity measure on all possible pairs. This produces adjacency matrices where connectivity estimate between signal i and j is the ijth (and jith, due to symmetry) entry. We chose to analyse the matrices obtained from both the coherence and the PLI [115] to look for differences between network topologies of zero and non-zero phase lag dependencies in the channels [30].

3.4 Results

3.4.1 Index Comparisons

Fig 3.3 shows the index curves (i.e. index plotted against network density) for C, V, Q, R, L and H (all introduced in Section 2.5 except R (3.1)) for all archetypes as well as for the EEG PLI (red shade) and coherence (blue shade) networks. From these plots we can note proxy maximal and minimal topologies for the given topological characteristics. These maximal and minimal topologies are explained as the curves whose lines are consistently lowest or highest over all densities. Fractal Modular networks (purple lines) are maximal for both C and Q (top left and centre left, respectively). This is to be expected since the modules are complete sub-networks with very few edges between modules, maximising Q. Further this restricts the number of open triples in the graph, maximising C, by restricting open triples to relating only to those few edges which do extend between modules. The star CWN (orange lines) acts as a maximal topology for V, as expected from the theory explained in Section 3.2. It is also a theoretical minimum topology for L (bottom left), since all node pairs which do not share an edge have a minimum path that is the next smallest possible– 2– traversing through any of the nodes of degree $n - 1$. Regular graphs, such as the ring lattice network (blue lines), give 0 degree variance and hierarchical complexity, thus are minimal topologies of these features.

Comparing the plots in Fig 3.3 of C (top left) with L (bottom left), and R (centre right) with H (bottom right), it is immediately clear that L and H show extreme behaviour at low densities while remaining consistent at higher densities. This exemplifies how these indices are aimed

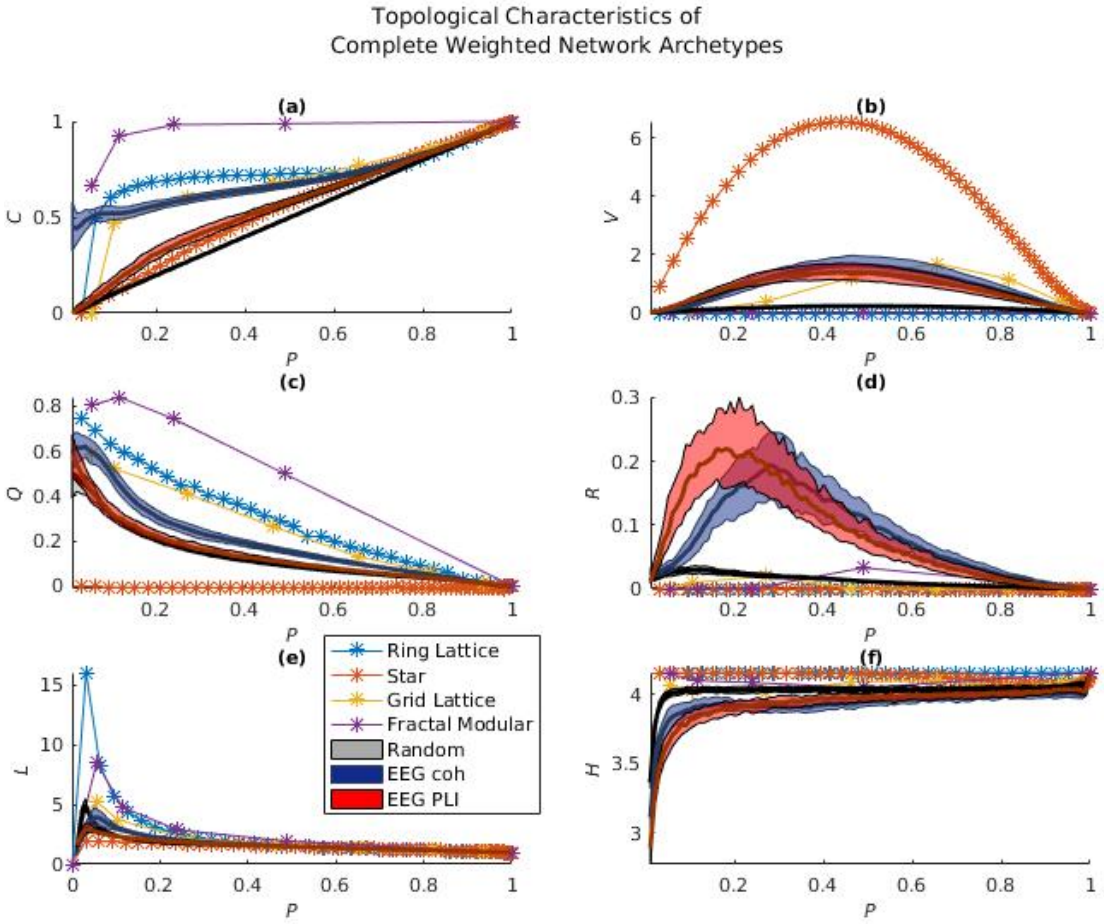

Figure 3.3: Topological index values for integration (C), regularity (V), modularity (Q), hierarchical complexity (R), characteristic path length (L) and network entropy (H) against network density, P. Curves relate to network models as indicated in the legend (bottom right). Random and EEG curves show the median, with interquartile range shaded, over all realisations/subjects.

at analysis of sparse networks, where it appears that values can take a much greater range than for higher density networks.

To explore these comparisons further we perform statistical analysis with population t-tests on the differences of distributions of index values of EEG PLI and E-R random networks as well as of EEG PLI and EEG coherence networks (Fig 3.4). The results show that C (right) and R (left) attain a greater range over edge density, P, of significant differences than their counterparts, L and H. Particularly, R distinguishes differences in 77 of the densities analysed with an average effect size (Cohen's d) of 0.6413 in the EEG PLI and coherence comparison (solid blue line), whilst H finds 71 differences with a average effect size of 0.5976 (solid

yellow line). Comparing the EEG PLI networks with E-R random networks (Fig 3.4, left, dashed lines), both indices find differences at all levels, but the mean effect size found by R, 1.4054, exceeds that of H, 1.2024. Thus, overall our index outperforms H in both magnitude and range of differences found.

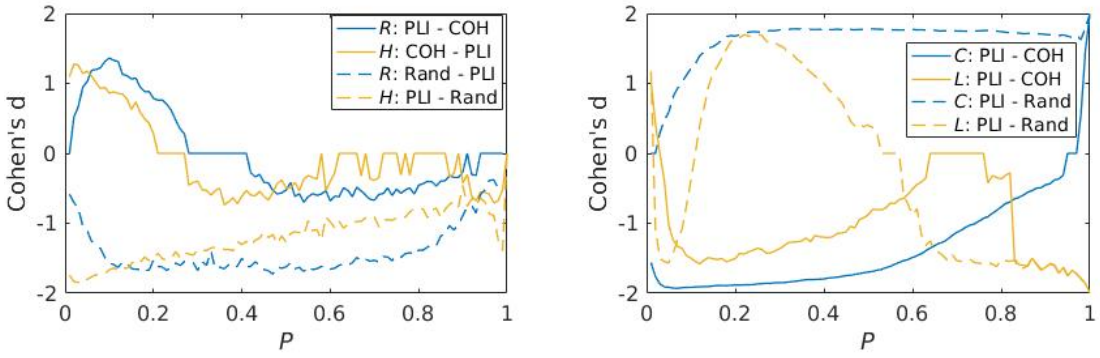

Figure 3.4: Effect sizes of significant differences found between PLI and coherence networks (solid lines) and PLI and random networks (dashed lines) using complexity indices R and H, left, and integration/segregation indices C and L, right. These are computed for, and plotted against, each integer percentage density, P. A zero indicates no significant difference found.

Similarly, C finds a greater range and magnitude of differences than L, Fig 3.4 right. For PLI vs coherence, C discerns differences at 97 densities with an average effect size of 1.4599, while L finds 86 differences with an average effect size of 1.1909. For PLI vs random networks, C discerns 98 differences with average effect size of 1.6502 and L discerns 94 differences with average effect size of 1.2979. Furthermore, L displays inverse differences at low densities (1-12%) compared to higher densities in the PLI vs random comparison (dashed yellow line). This inconsistency is undesirable for translatability of integrative behaviour of network types from sparse networks to more dense networks.

Given these results, for the rest of our analysis, we will drop L and H and focus on the four proposed indices, C, V, Q and R. We must emphasise that this is taken purely in terms of the simplicity of explaining a general topological factor and does not mean that L and H are not useful for other purposes.

3.4.2 Weighted Complex Hierarchy Null Model

Fig 3.5 shows the mean results of C (top left), V (top right), Q (bottom left) and R (bottom right) over 100 realisations of each of the WCH models. We include a reduced

number of strength parameters in the figure ($s = 0.1, 0.2, \ldots, 0.7$) than those computed ($s = 0.05, 0.1, \ldots, 0.75$) for greater clarity. Above 0.75 the parameter begins to saturate as the weights of the hierarchy levels tend to linear separability (linear separability occurs when $s = 1$ since $0s, 1s, 2s, ...$ then places the edge weights, originally in $[0, 1]$, in disjoint ranges $[0, 1], [1, 2], [2, 3], ...$). We see that WCH networks (grey shaded lines) exhibit curve behaviour similar to the EEG networks and E-R random graphs (as in Fig 3.3). The scale-free model (red error bars) also exhibits a similar behaviour, however in stark contrast, the small-world model (blue error bars) exhibits very different behaviours than those of the EEG or WCH networks, exhibiting a strong unsuitability for comparisons with EEG networks with much higher modularity and highly right skewed V curve (Fig 3.5, top left) towards high densities as well as a similar right skew in R (bottom left) which is opposite to the left skew found for WCH and EEG network types. Although the scale-free model exhibits similar tendencies in topological indices to the WCH and EEG networks, its range of values and densities is clearly very limited and so, therefore, its ability for topological refinement.

By increasing the strength parameter of the WCH model we change the topology in a smooth fashion with decreasing integration, regularity and modularity (Fig 3.5, top left, top right and bottom left, respectively). Interestingly, R (bottom right) rises with increasing strength parameter from $s = 0.05$ up to $s = 0.3$ where it takes its maximum values at densities ranging from 1-30% before falling again from $s = 0.35$ until $s = 0.7$. Further, above $s = 0.3$, the curves begin to deviate significantly from those of the EEG PLI networks, exhibiting greater plateaus of high complexity (lighter grey lines) which are more comparable with the EEG coherence networks.

Interestingly, the complexity of the EEG PLI networks appears to attain maximal values of R of all the networks studied here (Fig 3.3). The only model which comes close is the WCH model (Fig 3.7, bottom right). To clarify this observation we perform population t-tests on R values of the EEG PLI networks against that of the WCH model with strength parameters ranging from $s = 0.2$ up to $s = 0.4$, i.e. two steps before and after the maximal complexity setting of $s = 0.3$. The results are displayed in Fig 3.6. In the vast majority of instances of strength parameter and density, the EEG PLI networks do indeed exhibit greater complexity than the WCH model. Exceptions to this are shown between network densities in the range

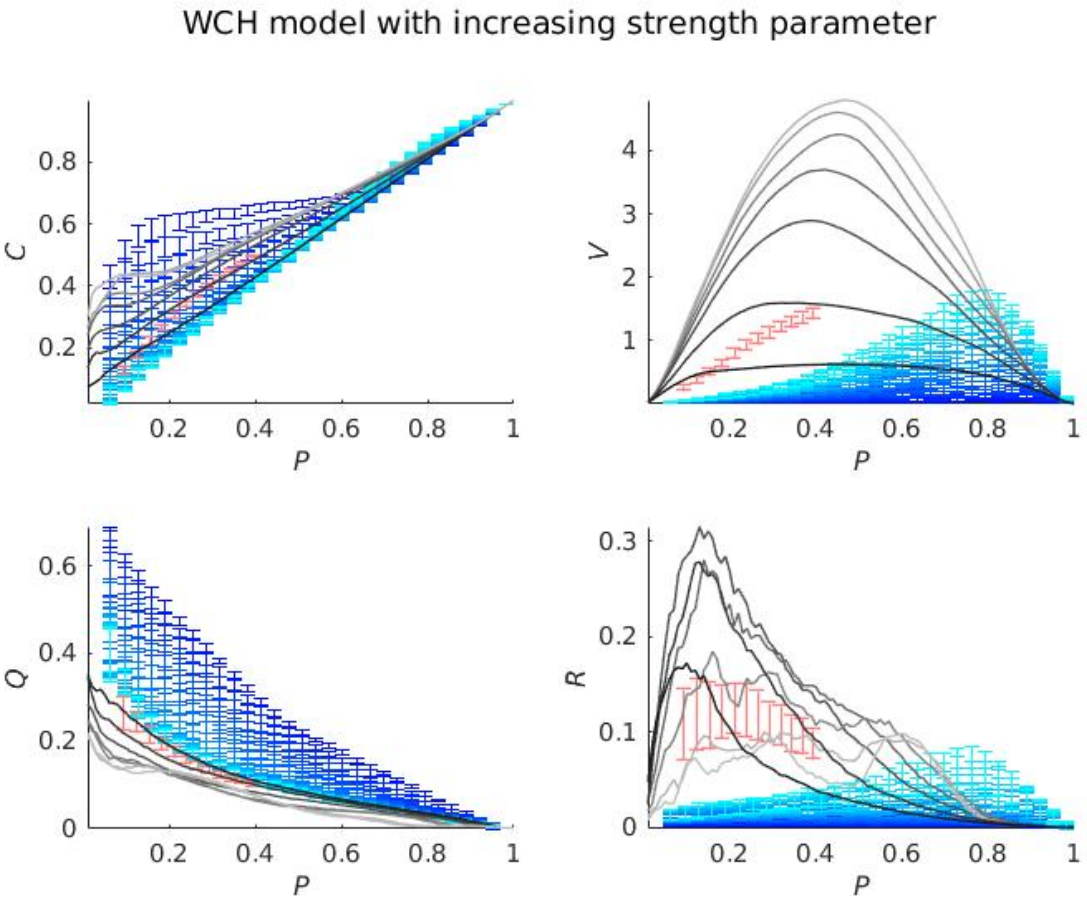

Figure 3.5: The topological characterisation of network models by clustering coefficient, C, degree variance, V, modularity, Q and hierarchical complexity, R, plotted against network density, P. Grey lines indicate mean values of the weighted complex hierarchy model with increasing light shade indicating increasing strength parameter from $s = 0.1$ in steps of 0.1 up to $s = 0.7$. Red error bars indicate values of the Albert Barabasi scale-free model. Blue errorbars indicate values of the Watts Strogatz small-world model with increasingly light blue indicating increasing proportion of edges being randomly rewired.

$0 - 0.2$, this is particularly strong for the WCH model with parameter $s = 0.3$ although effect sizes in this range only get just above 0.5, which is still quite low. Also, as the weight parameter increases, the high plateaus previously mentioned begin to take effect as in the medium ranges of density the R values of the PLI networks and WCH model becomes more indistinguishable, with greater complexity found in the range 40-70% in the WCH models with $s = 0.35$ and $s = 0.4$ with effect sizes reaching above 1.

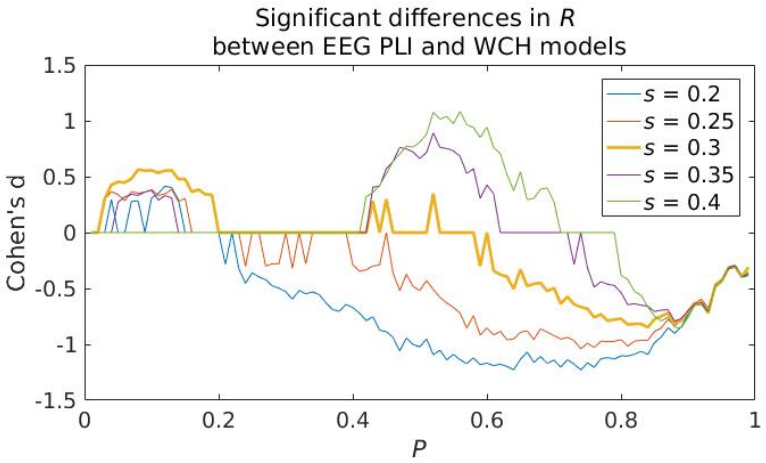

Figure 3.6: Effect sizes of significant differences found between EEG PLI networks and WCH models with parameters as indicated in the legend. Zero indicates p-value insignificant at 5% level. P is the network density.

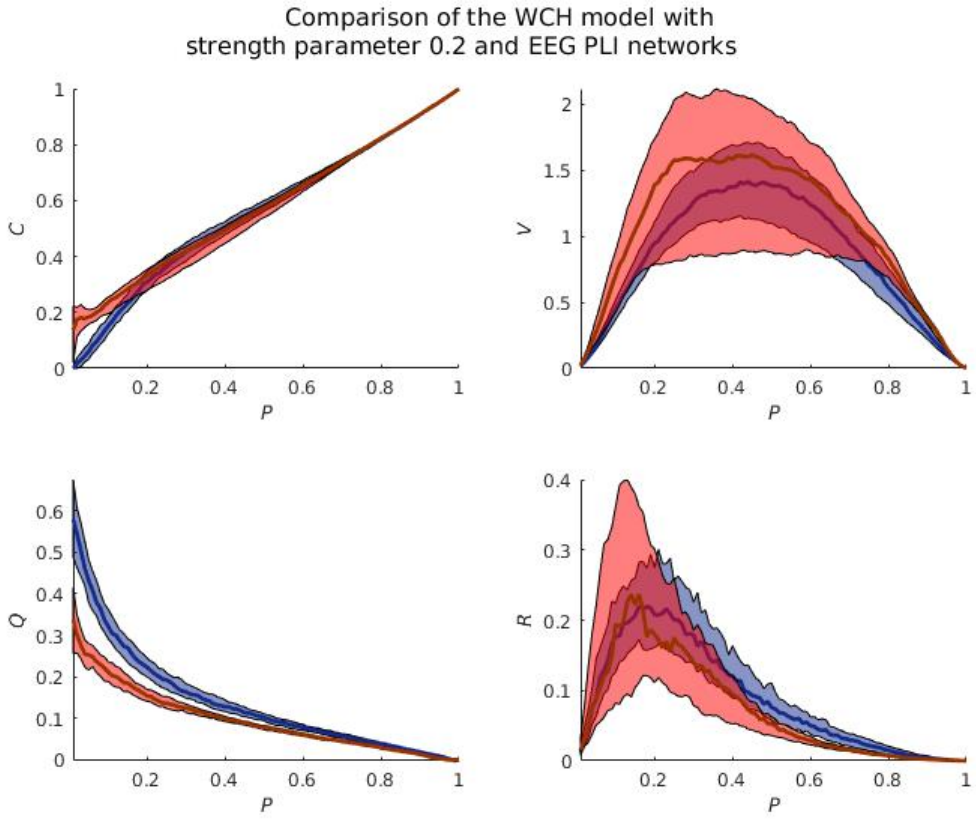

Figure 3.7: Clustering coefficient, C, degree variance, V, modularity, Q, and complexity, R, against edge density, P, of binarised weighted networks, for WCH model (red) and EEG PLI neworks (blue). Curves show median with interquartile range shaded.

3.4.3 Null model approaching EEG phase-lag networks

Fig 3.7 shows the values of the four topological features– hierarchical complexity, integration, regularity and modularity for EEG PLI networks and the WCH network with strength parameter 0.2. We see clearly that these networks behave very similarly with respect to the given indices. The most obvious difference is that the modularity, Q, of PLI EEG networks is higher (bottom left). Also, as previously discussed, the PLI network complexity is greater than the WCH model, but it is still by far the most comparable model for complexity of those presented here.

3.5 Discussion

The behaviour demonstrated by the WCH model with respect to R indicates that high complexity arises from a hierarchical structure in which a greater degree of variability is present in the rankings of weights with respect to hierarchy level. Too little difference between levels and the hierarchy is too weak to maintain complex interactions, too much difference between levels and the complexity of the hierarchy is dampened by a more ordered structure produced from the tendency towards linear separability of the edge weights enforced by the strength parameter. Thus, we provide evidence that topological complexity is not driven by integration, arising as a middle ground between regular and random systems as previously conjectured [52, 99], but, driven by hierarchical complexity, arising in the middle ground between weak hierarchical topology or 'all nodes are equal' systems, such as random or regular networks, and strong hierarchical topology, such as star or strict class-based systems including grid lattice and fractal modular networks (see Fig 3.5). Thus the hierarchical structure can be seen as a key aspect of the complexity inherent in complex systems.

The PLI EEG networks display a generally greater hierarchical complexity than that expressed by our model which is specifically designed to probe complex interactions in hierarchical structures. Thus we pose such complexity as a key aspect of brain function as modelled by phase-based connectivity.

There are two clear reasons why the WCH model is a good fit for functional connectivity networks from EEG recordings. Not only does it create several hub like nodes giving a high

degree variability, but furthermore it simulates the rich club phenomena found in complex brain networks [103, 116], as the higher the hierarchy levels of two nodes, the stronger the weight of the edge will be between them, see Fig 3.1.C. Specifically, the rich-club coefficient is defined as $\phi(k) = 2E_{>k}/N_{>k}(N_{>k} - 1)$, where $E_{>k}$ is the number of edges between nodes of degree greater than k and $N_{>k}$ is the number of nodes of degree greater than k. Now, let us consider the nodes in the top level of the hierarchy as a possible rich-club in a WCH model with l levels. Then the additional weight given to each such rich-club edge in the model is the maximum additional weight of the model– $2(l-1)s$. Relating this to probabilities, the edges with the most expectation to exist are precisely the "rich-club" edges and thus $\phi(k)$ will be large in our model and will increase with increased weight parameter s as the randomness provided by the ER weights, w_{ER}, becomes less prominent to the overall weight $w = w_{ER} + 2(l-1)s$.

One of the greatest benefits of this model over others is that it simulates brain networks previous to network processing steps because it creates CWNs rather than sparse networks. This means that any and all techniques one wants to use on the brain networks can be applied elegantly and in parallel with this single null model free from any complications. Particularly, methods which create sparse binary networks directly, whether these models are built independently from the brain networks [53, 56] or are constructed by the randomisation of edges of the networks being compared [45, 52], run into problems with density specification (in the case of independent models) and reproducibility (in both types of model). With the WCH model, we can simply create a bank of simulated CWNs which can be used throughout the study in exactly the same way as we use the functional connectivity CWNs.

As an example of the power and elegance of the proposed model, say we want to find maximum spanning trees [43] of our brain networks and compare with a null model, then we simply take the maximum spanning trees of our null model. In contrast, in [44] they use a convoluted reverse engineering process by assigning random weights to the edges of Watts-Strogatz small world networks (which are themselves of limited comparability to brain networks) and compute the MST from these resulting sparse weighted networks.

Further, as seen in Fig 3.1.C, for technical studies which rely on network simulations, the WCH model is built on parameters which can be altered to subtly change the resulting topology. This allows for sensitive analysis of a new techniques ability to distinguish subtle topological

differences. Such paradigms are evident in clinical studies where, for example, one may try to distinguish between healthy and ill patients [17,110] or between different cognitive tasks [117], so that this null model offers simulations which are directly relatable to clinical settings, as we shall explore further in chapter 4.

We see there is a large difference in the integration, modularity and complexity of the EEG coherence and PLI networks (Fig 3.3, top left, centre left and centre right, respectively). The EEG coherence networks (blue shade) behave similarly to the ring (blue lines) and grid lattice (yellow lines) networks, agreeing with the volume conduction effects that dominate zero-lag dependency measures [30], i.e. the closer the nodes are the stronger the weights are. The PLI networks (red shade) on the other hand have a more integrated and less modular nature, which reflects the notion that phase-based functionality mitigates volume conduction effects and is thus less confined by anatomical structure [30]. The very high complexity of the PLI networks (and very possibly phase-lag measures in general [29]) provides evidence to support that phase-based connectivity does indeed largely overcome the volume conduction effect and therefore maintains a richer complexity echoing the complex interactions of brain functionality [16].

With regards to how the WCH model advances our understanding of PLI and coherence network differences, we note that the high segregation of the coherence networks (Fig 3.3 top left) is approached by the WCH model with high values of strength parameter (Fig 3.5, top left) and is comparable with regular lattice and grid lattice CWN curves (Fig 3.3, top left, blue and yellow lines, respectively), denoting a move to a more strict class-based topology. This is also reflected in the hierarchical complexity (bottom right of corresponding figures), where the lower complexity peaking at a later density to PLI (Fig 3.3, centre right) is mimicked in the behaviour of increasing strength parameter in the WCH model (Fig 3.5, bottom right). This provides further evidence of the relevance and flexibility of the WCH model. In contrast there is an evident lack of ability to make similar comments with respect to the popular small world and scale-free models. This criticism can be extended towards network models which randomise edges while maintaining degree distributions [45], since such an enforced topological attribute does not allow one to analyse how that very important attribute is actually constructed.

A striking feature seen is in the degree variance curves where a highly symmetric parabolic

curve is noted with a central maximum value for random graphs, WCH networks and EEG networks. This feature reveals to us that the scale-free paradigm known in the degree distributions of sparse networks may be part of a broader picture of consistent behaviour at higher densities also which can be detailed using weighted networks. Noted in our results, as the density of the network increases one obtains more even distributions of high and low density nodes, indicated by the high values of V, and, eventually, towards high densities the symmetry of V values with low densities tells us that the scale-free network is characterised by a small number of low degree nodes and a majority of high degree nodes, i.e. the inverse (or complement) of the low density behaviour.

3.6 Conclusion

We introduced an index for measuring the hierarchical complexity of a network and a highly flexible and elegant WCH model. These provided key insights into what distinguishes functional brain networks from both ordered and spontaneous forms as generally the most complex kind of topology and the important role that hierarchical structure plays in this. Further, we showed that phase-based connectivity topology was more complex than amplitude influenced connectivity topology, which we extrapolated as due to the more ordered structure enforced by volume conduction effects. In our analysis we constructed a framework for CWNs for brain functional connectivity to replace the framework for sparse networks adopted from other network science research areas. This included the synthesis of concepts from the literature in a succinct manner and the generalisation of sparse binary archetypes to CWN form. The perspective allowed by this comprehensive analysis provided new evidence regarding key factors of network topology in general. For example, we note possible parallels of the scale-free paradigm for all network densities through binarisations of weighted networks. Particularly, these insights help towards a comprehensive understanding of the framework within which functional connectivity networks are set and thus provide invaluable information and tools for future clinical and technical research in neuroscience. Matlab codes for all synthesis and analysis of the networks as introduced in this chapter are publicly available at http://dx.doi.org/10.7488/ds/1520.

Chapter 4

Accounting for the Complex Hierarchical Topology of EEG Functional Connectivity in Network Binarisation

The contributions of this chapter were published in PLOS ONE in October, 2017 [2].

4.1 Introduction

Selecting a method to binarise functional connectivity networks is a major step in network construction in which the researcher is presented with a large degree of subjective choice [26, 27, 30, 40]. Because of this, recent research emphasises the importance of solutions to the binarisation problem in functional connectivity [30, 44, 47, 50, 111, 118–121]. While some find sparsity desirable based on the physiological hypothesis that function should be regarded as emerging through physically connected regions explained by a low wiring cost [38, 44, 50], in real data analysis others found higher densities to be as or more relevant [37, 49, 120]. We attempt to gain a foothold on answering why analysis of the data may not conform to the sparsity hypothesis by determining how the informational density of hierarchical structure of the functional network contributes to useful binarisation.

As shown in Chapter 3, the hierarchical structure of EEG functional connectivity is a key aspect of its informational complexity [1]. Also shown was that functional connectivity is characterised by high degree variance which is indicative of the large range of the general strength of network nodes. That is, one can expect that certain nodes have generally large adjacent weights, while others may have generally small adjacent weights. For a given node,

the relativity of weight magnitudes of edges adjacent to the other nodes in the network is thus important to keep track of throughout the network and not just in the largest connections and nodes in the highest hierarchy levels as would be promoted in sparse densities, see Fig 4.1. Thus we propose that any useful binarisation technique for functional connectivity should necessarily be able to account for the density of information inherent in a broad complex hierarchical structure. It should follow that sparsity is not necessarily a desirable feature of brain functional networks and also that statistical thresholds on a case by case basis of the connectivity computations does not necessarily translate to a topological advantage in the resulting networks.

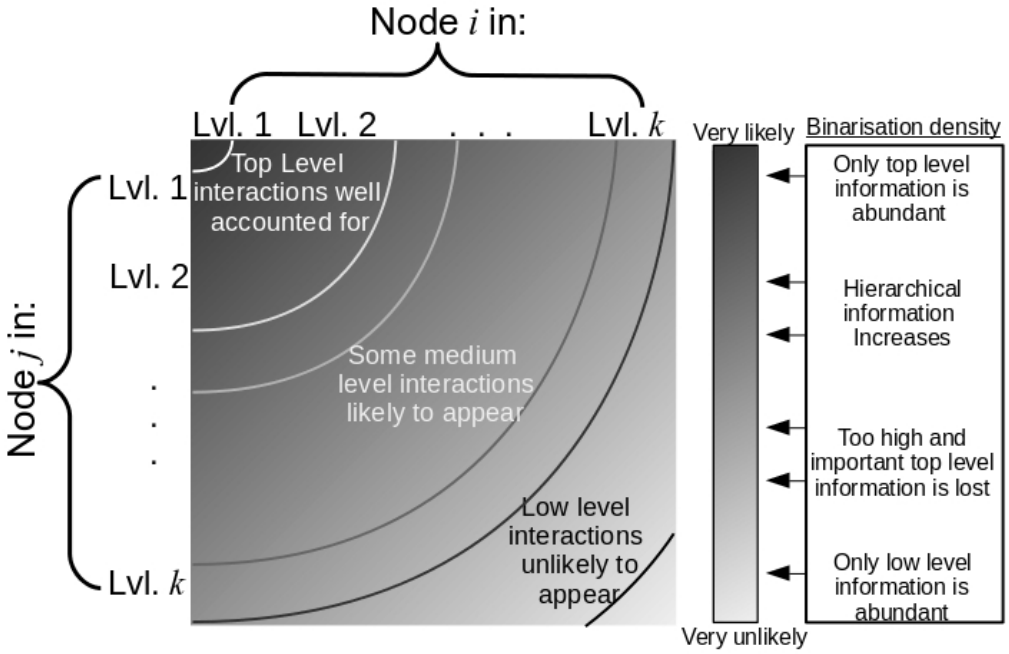

Figure 4.1: Illustration of the likelihood of an edge appearing between nodes i and j, in hierarchy levels denoted by the x and y axes, in the binarised form of a weighted hierarchical network. Left, the effects of increasing binarised network density (strongest weights kept) on the hierarchical information of the network where black indicates 0% density and white indicates 100% density.

To provide rigorous simulation results for binarisation techniques, we implement the WCH model [1]. Since the parameters of the model provide a fine-tuning of hierarchical topology, we can create a ground truth of subtly different topologies. This can be exploited to assess the ability of binarisation methods to correctly identify topological differences between networks. Here, we intentionally do this to echo the set-up of a neuroscientific study and make the

simulations as relevant as possible to the research community. This follows since, for the lack of a ground truth, studies in network neuroscience are generally based on contrasting conditions such as in cognitive tasks or contrasting populations where, for example, network features of patients are compared against those of healthy, age-matched controls [17].

Using simulations, we seek to clarify how network size and density range may affect the ability to discern small topological differences in network topology. In analysis, we compare state-of-the-art non-arbitrary binarisation techniques– MST [43, 44], USP [47], CST [42, 117] and ECO [50]– with a number of arbitrary percentage thresholds, as well as the original weighted networks to directly compare binary and weighted approaches.

We then analyse these techniques when the simulations are subject to random and targeted topological attacks. We regard these as random and targeted attacks [122] which preserve the network size. This is desirable given that many index values are dependent on network size [123]. By randomising a percentage of weights in populations of subtly different complex hierarchical networks in parallel we can test how well the binarisation techniques can still uncover the differences between these populations under varying sizes of attack. We implement these analyses to test the binarisation techniques' robustness in representing true network characteristics in the face of noise and/or outliers in the estimation of coupling between brain time series.

We go on to apply our non-arbitrary binarisation techniques to three real EEG datasets. We compare our thresholds on distinguishing the well known Alpha activity existing between eyes open vs eyes closed resting state conditions in healthy volunteers with a 129 channel EEG [124]. We then compare these techniques for distinguishing VSTM binding tasks in healthy young volunteers with a 30 channel EEG [117]. Finally, we compare our techniques in distinguishing between AD patients and healthy control in a 16 channel EEG set-up [125]. The varying sizes of these networks provides evidence for the translatability of the methods to different network sizes in the applied setting. The scripts, functions and data sets used in this study are available at the University of Edinburgh's data depository: http://datashare.is.ed.ac.uk/handle/10283/2783.

4.2 Methods

This section details the network simulations (4.2.1), binarisation techniques (4.2.2), network indices (4.2.3) and statistical tests (4.2.4) used in this study.

4.2.1 Simulated experiments

4.2.1.1 Experimental design

For the simulations, we follow the procedure as illustrated in Fig 4.2. WCH models are generated as a ground truth to test the ability of binarisation techniques to distinguish subtly different populations of size 20. These different populations are generated using realisations of the WCH model with small differences– 0.05– in the strength parameter, s. The procedure thus follows that of a typical clinical study, where small populations of contrasted conditions are analysed using network science techniques with statistical tests used to determine significance of the differences between the populations. This methodology is used for networks with 16, 32, 64 and 128 nodes, spanning a large range of network sizes as used in current research, e.g. see [17].

4.2.1.2 Random and targeted topological attacks

We test the robustness of the given binarisation techniques by subjecting these same simulated networks to random and targeted topological attacks before implementing similar topological comparative analysis as above. Random and targeted attacks were originally formulated by deleting entire nodes from the network [122]. We implement a weight randomising approach, thus preserving network size which is important when comparing different indices [123]. Further, this is more relevant to brain networks where the network size is determined a priori, but rather the information recorded at the nodes are susceptible to attacks such as signal noise or corruption. These random topological attacks are implemented by substituting the WCH model's weights with corresponding non-zero entries of a sparse, uniformly random and symmetric weighted adjacency matrix. The steps involved here are i) generate a sparse random weighted matrix, $\mathbf{E}$, using the sprand function in Matlab with a particular density and the same

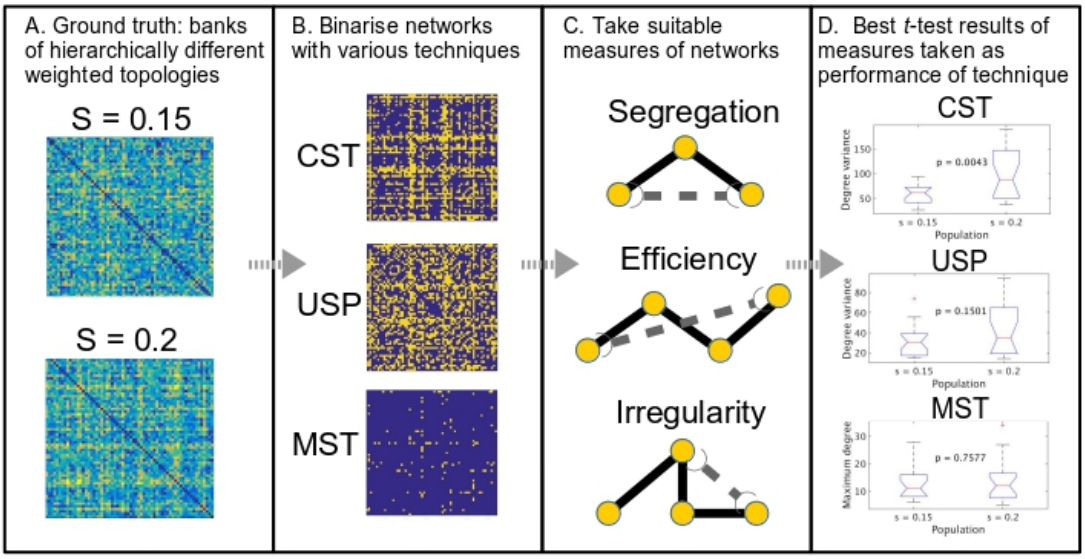

Figure 4.2: Methodological steps in the evaluation of binarisation techniques for determining ground truth topological differences.

size as the adjacency matrix of interest, $\mathbf{W}$. ii) set the diagonal of $\mathbf{E}$ to zero to ensure no loops. iii) make $\mathbf{E}$ symmetric by mirroring the lower triangle to the upper triangle of matrix. iv) replace all w_{ij} with e_{ij} for $e_{ij} \neq 0 \in \mathbf{E}$. We implement this comparison by increasing the density of the sparse matrix, i.e. densities of $0, 0.05, 0.1, ..., 0.95, 1$. Targeted topological attacks are implemented similarly except the attacks are restricted only to those nodes whose average adjacent weight is over one standard deviation above the mean, relating to those nodes with abnormally strong connectivity. Such strongly weighted nodes are known as hub nodes for their importance to the topology of the network.

4.2.2 Network binarisation

Details of the topological binarisation methods used are, other than as detailed below, found in Section 2.4.

4.2.2.1 Cluster-span threshold

The CST chooses the binary network at the point where open to closed triples are balanced [49]. This balance occurs when $C = 0.5$, which is obvious from the definition. The algorithm for the

CST computes the binary networks for each possible number of strongest edges between 15% to 85%, rounded to the closest real edge density. The clustering coefficient is then computed for each of these networks, obtaining a vector $\mathbf{C} = \{C_{15}, C_{16}, \ldots, C_{84}, C_{85}\}$, where C_i is the clustering coefficient of the binary network at the ith % density, rounded to the nearest number of edges. Then the network of the CST is the binary network corresponding to $Z = \text{argmin}_i(C_i - 0.5)$, i.e. the threshold achieving minimum value of the vector $\mathbf{C}$ minus the clustering coefficient value which obtains an equilibrium between triangles and non-triangle triples, 0.5. The values of 15% and 85% are chosen as safe values based on experimental evidence [42, 49]. Particularly, below 15%, real brain networks can have a tendency to fracture into more than one component, thus making calculations of metric values, including the clustering coefficient, inconsistent and unreliable [37].

4.2.2.2 Union of Shortest Paths

For the union of shortest paths we will use weight transformation $\hat{\mathbf{W}} = -\ln(\mathbf{W})/\alpha$, where

$$\alpha = \min\{\mathbb{N}\} \text{ s.t. } \max_{i,j}(\hat{w}_{ij}) < 1, \tag{4.1}$$

as it has been shown to offer a better spread of metric magnitudes which is important for shortest path problems [42].

4.2.2.3 Arbitrary proportional thresholds

Arbitrary thresholds can be implemented by either choosing a weight above which edges are kept and below which edges are discarded, or by choosing a percentage of strongest weighted edges to keep in the network. The latter choice is more robust and easier to compare between different set-ups and subjects because it keeps the connection density constant and thus is not affected by the values of the weights, which may vary wildly particularly when considering the comparison of different connectivity measures. In order to cover the density ranges of both the sparsity hypothesis and the hierarchical complexity hypothesis, we choose arbitrary thresholds which maintain the strongest 5%, 10%, 20%, 30%, 40%, and 50% of edges to make sure we cover the relevant array of connection densities whilst reducing redundancy. Note, very sparse densities are already covered by the MST and ECO thresholds.

4.2.3 Network indices

To analyse the simulated and real EEG networks we use a variety of common metrics. As well as L, E, D, C, P and V in Chapter 2, the following indices for weighted networks and MSTs will be used:

- The **diameter**, D, of a graph is the largest shortest path length between any two nodes in the graph [44].
- The **weighted clustering coefficient**, $C_{\mathbf{W}}$, is a weighted version of the clustering coefficient for binary networks [39].
- The **leaf fraction**, *LF*, of a tree is the fraction of nodes in the graph with degree one. Note, every path containing such a node either begins or ends at that node [44].
- The **edge density**, P, is the ratio of the number of edges in the graph to the total possible number of edges for a graph with the same number of nodes, i.e. $P = 2m/n(n-1)$ [6]. For the CST, P takes an inversely relational position to C of proportional thresholds. This can be seen by considering two weighted networks whose values of C increase monotonically with increasing P and such that one has higher values of C than the other, which is a working assumption in our case. Then the network with the greater values of C will attain its CST at a lower density, P. In a similar vein, P of the USP is inversely related to L of proportional thresholds– the higher the density of the USP, the shorter the average shortest path in the weighted network.
- The **maximum degree**, *MD*, of a network is just as named– the degree of the node with the most adjacent edges in the network [44].

Note that, though we introduced R in Chapter 3, we do not use it here since, as yet, there is no weighted equivalent and we wish to be as fair as possible in our comparisons using well documented and widely used indices.

For each binarisation technique we choose three metrics to analyse the subsequent binary networks. These differ for each technique because of the construction of the network. Particularly, the MST metrics are chosen based on the study of Tewarie *et al.* [44]. Similarly, we choose three weighted metrics for analysing the original weighted networks. These choices can be found in Table 4.1. We try as much as possible to stick to three main categories of

metrics for each binarisation technique: segregation (M1 in Table 4.1), efficiency (M2) and irregularity (M3) [1, 16]. This notably deviates for M3 in the weighted case where the mean weight of the network edges, $\mu_{\mathbf{W}}$, is an appropriate and more obvious choice of metric than the variance of those weights.

Table 4.1: Grouped Topological Indices- Three for Each Network Type.

Index	CST	MST	USP	ECO	Weight	%T
M1	P	LF	C	C	$C_{\mathbf{W}}$	C
M2	L	D	P	L	E	L
M3	V	MD	V	V	$\mu_{\mathbf{W}}$	V

4.2.3.1 Functional connectivity

For the real EEG datasets, FIR bandpass filters were implemented for Alpha, Beta or both bands, as specified in the Materials, Section 4.3. A filter order of 70 is used to provide a good trade-off between sharp transitions between the pass and stop bands while keeping the filter order low in comparison with the length of the signals. The filtered signals were then analysed for pairwise connectivity using the PLI [31].

4.2.4 Statistical testing

Each simulation iteration undergoes 50 simulated trials of two populations of 20 networks. A population of networks is selected from a bank of 1000 WCH networks with given strength parameter. These banks exist for $s = 0,\ 0.05,\ 0.1,\ 0.15,\ 0.2,\ 0.25$ and 0.3. The other population in the trial then comes from a WCH bank with strength parameter with 0.05 difference. We undergo such trials for all possible combinations of 0.05 differences. We binarise these networks using each of our binarisation methods. We then compute the three indices, M1, M2 and M3, for each of these networks (weighted indices are computed from the original weighted networks). We perform population t-tests of these indices for the WCH binarisations. We choose the best index of the three to represent the ability of the binarisation method to discern subtle topological differences where the 'best' index is chosen as that which attains the maximum number of significant p-values out of the 50 simulated trials which are less than the standard $\alpha = 0.05$ level. If two or more indices obtain the maximum value, we then choose the one with the lowest mean log of p-values. Choosing the log in this instance

emphasises the importance of smaller p-values for distinguishing differences. The number of differences discovered, taken as a percentage of the total number of trials run, then represents the 'accuracy' of the binarisation technique at distinguishing the ground truth, i.e. that the topologies of the populations are different.

4.3 Materials

4.3.1 Eyes open - eyes closed resting state data

The eyes-open, eyes closed 129 node dataset is available online under an Open Database License. We obtained the dataset from the Neurophysiological Biomarker Toolbox tutorial [124]. It consists of 16 volunteers and is down-sampled to 200Hz. We used the clean dataset which we re-referenced to an average reference before further analysis. The data were filtered in Alpha, according to known effects [126], using an order 70 FIR bandpass filter with hamming windows at 0.5Hz resolution. For each subject we take one arbitrarily long 1000 sample epoch (5s) from the 1000-2000th samples.

4.3.2 Visual short-term memory binding task data

The stimuli were non-nameable shapes and non-primary colours known to be difficult to rehearse verbally [96, 127]. Two arrays of three items each were presented to the left and to the right of a fixation cross centred on the screen on a grey background (Fig. 4.3). Each array was presented in a virtual 3×3 grid, 4° horizontally centred, 8° vertically centred and 3° to the left and right from fixation. Each item took up 1° and the distance between items was never less than 2°. Items for the study display were randomly selected from a set of eight polygons and eight colours [96] and randomly allocated to 3 of the 9 positions within the grid. For the test display the items were randomly shuffled within the same locations used in the study display. Hence, items were never presented in the same locations across study and test displays, rendering location uninformative.

Trials were self-initiated. A fixation cross appeared in the centre of the screen and remained on throughout the trial. After a button press, 500 ms lapsed before the arrow clues were presented. Two arrows appeared for 200 ms one above and one below fixation which indicated which

of the two visual arrays (left or right) the participants were to attend. An interval of random duration selected from 400, 500 or 600 ms followed the cues. The study display was then shown for 200 ms. After an unfilled retention interval of 900 ms the test display appeared and remained visible until the participant responded.

In the Shape condition each array of the study display presented three black shapes. The test display also showed three shapes. In 50% of the trials the content of the test display matched the content of the study display ("same trial"). The test display for the "different trials" showed two new shapes. In the Bind condition each visual array consisted of three shapes in different colours. In the test display for the "different trials" two coloured shapes swapped their colours. The participants responded "same" or "different" by pressing two keys previously allocated with both hands. The participants completed 8 practice trials before undergoing 170 test trials for each of the conditions.

Each participant undertook four different conditions of the VSTM task which are distinguished by two different binary manipulations: 1. Shape or Bind: the test items consist of black shapes (shapes only) or shapes with colours (shape-colour binding). 2. Left or Right: the test items are shown on the left side or the right side of the screen (or hemifield) to which the participant is prompted before stimulus onset. The task was to detect whether or not a change occurred across two sequential arrays shown on an initial study display and a subsequent test display.

EEG signals were recorded for 23 healthy young volunteers while they performed VSTM tasks. These four tasks are categorised by two binary conditions- shape or binding and objects displayed in left hemifield or right hemifield, as described in Fig.4.3. For further specific task details, see [128]. Five of the volunteers were left-handed and eight were women. The mean and standard deviation of the age of participants and number of years of education is 23.0 ± 4.3 and 17.1 ± 2.8, respectively. Informed consent was obtained from all subjects. The study was approved by the Psychology Research Ethics Committee, University of Edinburgh, and methods in data collection were carried out in accordance with their guidelines.

The EEG data was collected using NeuroScan version 4.3. The EEG was sampled at 250 Hz. A bandpass filter of 0.01-40 Hz was used. Thirty EEG channels, corrected for ocular artefacts using ICA, were recorded relying on the 10/20 international system.

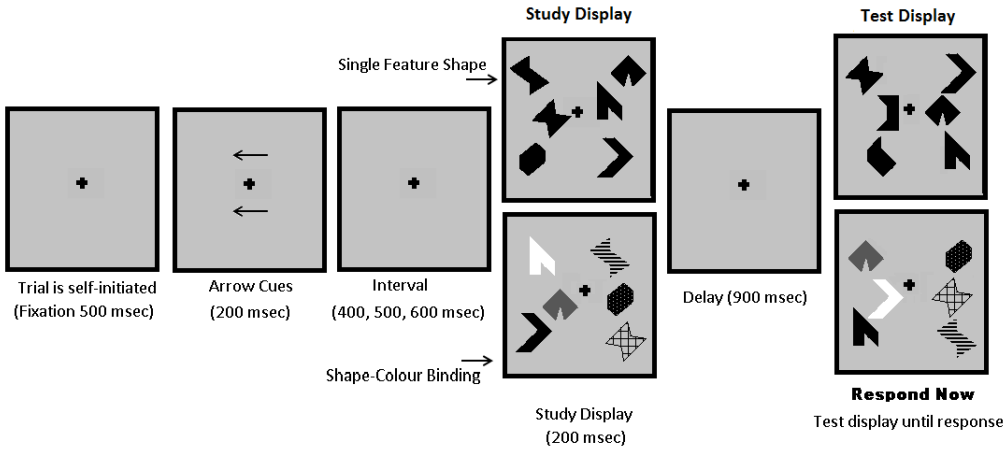

Figure 4.3: Chronology and design of the Visual Short Term Memory tasks. Arrow cues pre-stimulus indicate to the participant the hemifield being tested. The study stimulus consists of either black shapes (middle, top box) or shapes with associated colours (middle, bottom box). After a delay, the participant is presented with a test stimulus to discern whether the set of objects in the relevant hemifield are the same or different to the study stimulus. Courtesy of Dr. Mario A. Parra.

Further artefact rejection was conducted, rejecting trials which contained magnitudes of voltage fluctuations above 200 microvolts, transients above 100 microvolts and electro-oculogram activity above 70 microvolts. Only the trials with correct responses were kept as incorrect responses do not inform on working memory load in task comparisons. It is important to emphasise the distinction between a study of healthy brain function of task performance, as conducted here, for which the number of correct trials is not indicative, and a study of the performance of tasks by healthy people, for which the number of correct trials is indicative. In a few cases, no useful data was available for a volunteer performing one of the conditions resulting in an unequal number of volunteers per condition.

To keep comparisons straightforward, we chose only to look at those 19 participants of the original 23 for whom data on all of the conditions was available. We focused on the encoding (i.e., study display) and maintenance (delay) periods of VSTM, since these seem to be the stages of memory informing about the functional principles of organisation with regard to capacity and format of representation (Shape vs Bind) [127].

The mean $\pm$ standard deviation over participants for the number of kept trials for each condition were as follows: Shape, Left hemifield- 69.74 $\pm$ 6.67; Bind, Left hemifield- 63.79 $\pm$ 8.72;

Shape, Right hemifield- 66.32 $\pm$ 15.06; Bind, Right hemifield- 63.58 $\pm$ 16.26.

Based on relevant results [49, 117], the data were filtered in Beta using an order 70 FIR bandpass filter with hamming windows at 0.5Hz resolution. The epochs are 1s long and the number of trials is 65.7 $\pm$ 9.27 (mean $\pm$ SD). PLI adjacency matrices are averaged over trials. This dataset was provided by Dr. Mario A. Parra, Heriot-Watt University.

4.3.3 Alzheimer's disease data

The EEG recordings were taken from 12 AD patients and 11 healthy control subjects. The patients –5 men and 7 women; age = 72.8 $\pm$ 8.0 years, mean $\pm$ SD– were recruited from the Alzheimer's Patients' Relatives Association of Valladolid (AFAVA). They all fulfilled the criteria for probable AD. EEG activity was recorded at the University Hospital of Valladolid (Spain) after the patients had undergone clinical evaluation including clinical history, neurological and physical examinations, brain scans and a Mini Mental State Examination (MMSE) to assess their cognitive ability [129]. The ethics committee of the Hospital Clinico Universitario de Valladolid approved the study and control subjects and all caregivers of the patients gave their written informed consent for participation. The 16 channel EEG recordings were made using Profile Study Room 2.3.411 EEG equipment (Oxford Instruments) in accordance with the international 10-20 system. Full details can be found in [125]. The data were filtered both in Alpha and Beta as in [42], for separate analysis, using an order 70 FIR bandpass filter with hamming windows at 0.5Hz resolution. Recordings were visually inspected by a specialist physician who selected epochs with minimal artefactual activity of 5s (1280 points) from the data for further analysis. The average number of these epochs per electrode per subject was 28.8 $\pm$ 15.5 (mean$\pm$SD). This dataset was provided by Dr. Daniel Abasolo, University of Surrey.

The EEG PLI adjacency matrices used in this study are available at the University of Edinburgh's data depository at http://datashare.is.ed.ac.uk/handle/10283/2783.

Table 4.2: Grand average percentage of topological differences discovered between WCH models. Underlined are the best non-arbitrary technique (first rows) and any arbitrary density threshold that does as well or better than non-arbitrary techniques (last rows).

CST	**MST**	**USP**	**ECO**	**Weighted**
<u>71.3%</u>	22.4%	50.0%	53.4%	40.5%
T50%	**T40%**	**T30%**	**T20%**	**T10%**
70.4%	<u>71.5%</u>	67.3%	60.0%	41.3%

4.4 Results

4.4.1 Sensitivity to subtle topological differences in synthetic EEG connectivity

Fig 4.4 shows the results for differences discovered between WCH models with differences in strength parameter of 0.05. The grand averages are shown in Table 4.2. The CST is shown to outperform other non-arbitrary methods in general. In testing the comparisons of the WCH model with varying strength parameter, s, it discovers significant differences at the $\alpha = 5\%$ level 71.3% of the time over all strength comparisons and network sizes, as seen in Table 4.2. On the other hand, the MST discovers just 22.4% of the differences, the USP discovers 50% of the differences, the ECO 53.4 % and the weighted metrics discover just 40.5% of the differences. Out of these methods, in fact, it discovers the most differences in all but two cases– those being the 0.1 vs 0.15 comparison in the 16 node networks and the 0.25 vs 0.3 node comparison in the 128 node cases, of which the USP is best on both occasions.

In comparison with arbitrary percentage thresholds the CST appears to perform approximately the same as the 40% proportional threshold which discerns a slightly higher rate of 71.5% of differences over all cases. The 50% threshold also appears to be good at discerning differences here with an overall rate of 70.4% of differences discovered. These results agree with the hypothesis that complex hierarchical structures are best captured by larger density ranges. It is important to recall that we need non-arbitrary solutions rather than simply to find the best possible threshold for these specific simulations due to the possibility of over-fitting. With this in mind we can see that, compared to the other techniques, the CST outperforms the field in this study.

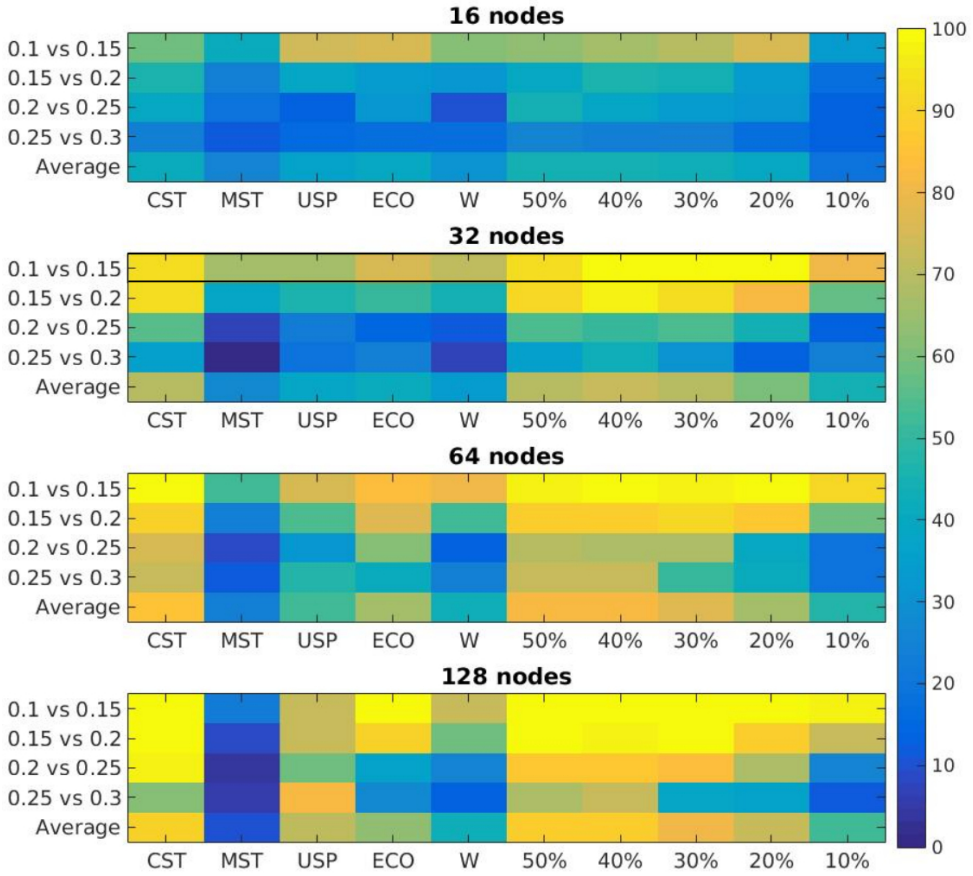

Figure 4.4: Percentage of topological differences discovered from population t-tests between WCH models. The y-axis shows $s = a$ vs. $s = b$ for WCH populations with strength parameter, s. The x-axis shows the binarisation method used where W is the weighted approach and percentages indicate arbitrary density thresholds.

4.4.2 Robustness to random and targeted topological attacks

The robustness to random and targeted topological attacks is evaluated by comparing the metrics from the attacked WCH models over all non-arbitrary binarisation techniques using population t-tests as before. For these analyses we look at the case in Fig 4.4 with the maximum ratio between the mean and standard deviation of accuracy over binarisation techniques, i.e. the case which maximises the ratio of average performance and comparability of performances across the compared methods. This ensures a level playing field for the methods where they all are performing fairly well. This happens in the 32 node case (6.3538) with differences in

strength parameter of $s = 0.1$ and $s = 0.15$, see the boxed row in Fig 4.4. The grand percentage over all sizes of attack for p-values below 0.05 for each metric is presented in Fig 4.5 for both random and targeted topological attacks. Generally, the binarisation methods as well as the weighted metrics are more robust to targeted attacks than to random attacks. Notably, the CST maintains the highest average accuracy of distinguishing topological differences for all metrics and both random and targeted attacks. This is particularly evident in the targeted attacks. For both kinds of attack, the weighted metrics come in second best while the ECO, the USP and MST perform relatively poorly.

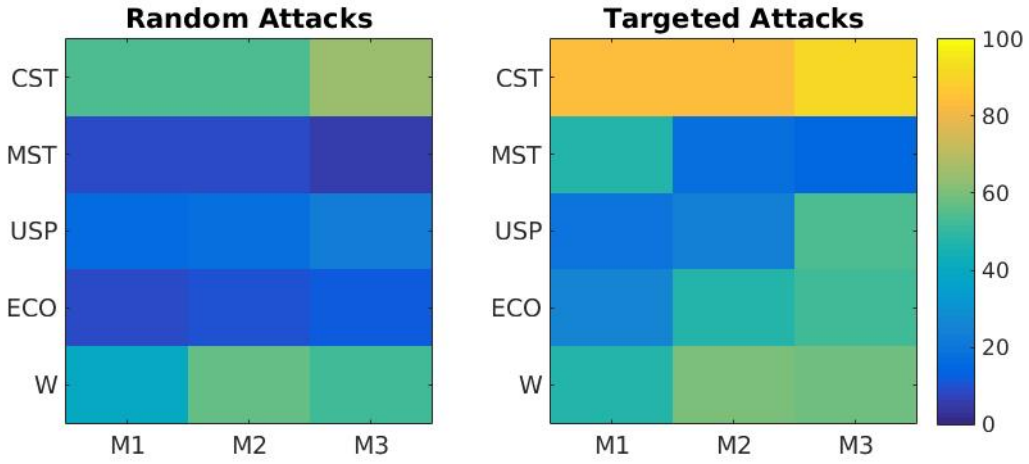

Figure 4.5: Percentage of topological differences discovered from population t-tests via binarisation methods (CST,MST, USP, ECO and weighted network (W)) between WCH models ($s = 0.1$ vs. $s = 0.15$) with random and targeted network attacks. Averaged over all sizes of attack. M1, M2 and M3 as in Table 4.1

The metric achieving highest accuracy for each binarisation technique and for each size of attack is shown in Fig 4.6 for both random and targeted topological attacks. Strictly in terms of robustness (with respect to depreciation from starting value) as opposed to best accuracy, the weighted networks prove the best with the least detriment noted by increasing the size of attack in its topological accuracy (green). The CST also does well here. For the random topological attacks, even at 50% of connections attacked, the CST notes an accuracy of 70% (blue line). The USP (yellow), MST (red) and ECO (purple) networks are not at all robust to random topological attacks in this scenario with immediate drop offs on the implementation of attacks.

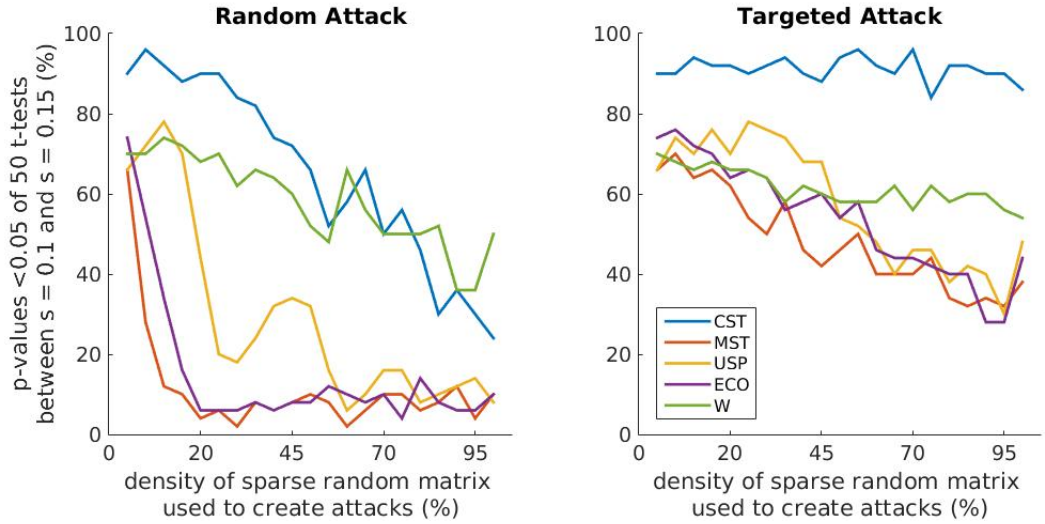

Figure 4.6: Percentage accuracy of method for distinguishing topological differences between attacked WCH models against the size of those attacks for random topological attacks (left) and targeted topological attacks (right). The values plotted are the maximum from the three indices, M1, M2 and M3 (as in Table 4.1), for the corresponding technique as indicated in the legend.

For the targeted topological attacks (Fig 4.6, right), the CST network (blue) shows the most resilience with no noticeable depreciation of accuracy. The other methods, in contrast, show a notable decrease in accuracy as more weights are randomised.

4.4.3 Real dataset results

We maintain our focus on comparing non-arbitrary methods since arbitrary approaches are inappropriate for neurophysiological studies where one can pick from an order of $n(n-1)/2$ thresholds.

Table 4.3 shows the results for distinguishing the difference in Alpha activity well known to exist between eyes-closed and eyes-open conditions [126]. The CST finds a significant difference in V of eyes open and eyes closed resting state activity indicating that the phase-dependent topology of EEG activity has lower hierarchical spread in the eyes-open condition implying greater hub dominance in the eyes-closed condition, see Fig 4.7, left. All of the weighted metrics also find significant differences. Neither the MST nor USP find any differences between these conditions. Probing further, $\rho(M1, M2)$ being the correlation coefficient of metric values across subjects of metrics M1 and M2 as defined in Table 4.1, the

weighted metrics in this case are all very highly correlated (all > 0.95 Pearson's correlation coefficient, ρ, Table 4.4) within condition. Therefore they cannot be seen to provide any distinct topological information. The corresponding correlations of the CST show a more distinct topological characterisation, also in Table 4.4.

Table 4.3: The p-values from paired t-tests between eyes open (EO) and eyes closed (EC) in Alpha band 129-channel EEG PLI Networks. Underline: Best value for each method. Bold: Significant values. M1, M2 and M3 as in Table 4.1

Index	CST	MST	USP	ECO	Weighted
M1	0.7504	0.4178	0.5063	0.6034	**0.0016**
M2	0.9319	0.4513	0.9942	0.5281	**0.0034**
M3	**0.0006**	0.9616	0.6577	0.6805	**0.0016**

Table 4.4: Pearson correlation coefficients of indices in Eyes Open (EO) - Eyes Closed (EC) Dataset for the CST and Weighted indices (wgt)

Index corr.	EC (CST)	EO (CST)	EC (wgt)	EO (wgt)
ρ(M1,M2)	0.7662	0.9692	0.9993	0.9512
ρ(M1,M3)	-0.0458	-0.7301	0.9999	0.9996
ρ(M2,M3)	-0.1695	0.6677	0.9994	0.9576

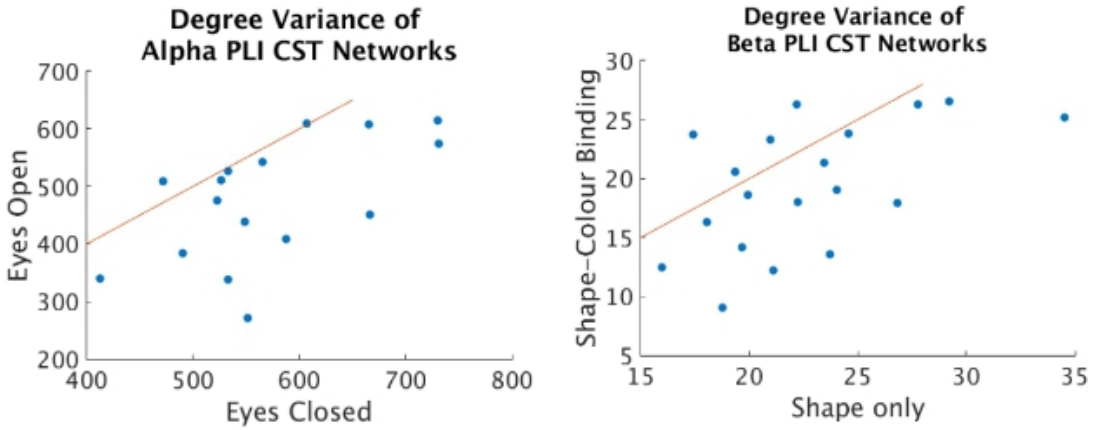

Figure 4.7: Scatter plots of degree variance for CST networks of Eyes Open vs Eyes closed resting state conditions in Alpha, left, and degree variance for CST networks of shape vs binding conditions in the Right screen in Beta, right.

Table 4.5 shows the results for distinguishing the difference in Beta activity existing between shape and binding tasks when tested in the Left and Right sides of the screen separately. A significant difference is found in V of the CST networks in the Right condition. This indicates

that the phase-dependent topology of EEG activity has less hierarchical spread in the binding condition implying greater hub dominance in the shape condition, see Fig 4.7, right. On the other hand, a significant difference ($p = 0.0059$) is found in C for the ECO networks in the Left condition. This indicates sparse density topology of EEG activity is less integrated in the binding condition.

Table 4.5: The p-values from paired t-tests between shape and binding tasks in Beta band 30-channel EEG PLI networks. Formatting as in Table 4.3

Hemifield	Index	CST	MST	USP	ECO	Weighted
Left	M1	0.5128	0.7186	-	**0.0059**	0.1007
	M2	0.0898	0.1383	-	0.1870	0.1010
	M3	0.8997	0.0911	-	**0.0238**	0.1010
Right	M1	0.5877	0.1919	-	0.5504	0.7742
	M2	0.9196	0.5716	-	0.5038	0.7733
	M3	**0.0088**	0.8146	-	0.2138	0.7733

Noticeably, the USP failed to find meaningful network information in this task because, even after transformation, all the weight magnitudes were in a range such that the shortest weighted path between each pair of nodes was the weight of the single edge joining them.

Table 4.6 shows the results for distinguishing the difference in both Alpha and Beta activity existing between AD patients and healthy age matched control. For the CST, an effect is noticed in V of Alpha activity (Fig 4.8, right) and a larger effect is found in the P of Beta activity (Fig 4.8, left). Since P of CST networks is inversely relational to C of arbitrary threshold networks, this tells us that Beta of AD patients is less segregated than control. Contrasting with this, the activity in Alpha suggest a more heterogeneous network in the Alpha band of AD patients than in age matched control.

Table 4.6: The p-values from population t-tests of network measures AD and control in 16-channel EEG PLI Networks

Freq. band	Index	CST	MST	USP	ECO	Weight
Alpha	M1	0.0852	0.3468	0.1167	0.7496	0.6736
	M2	0.3634	0.2630	0.1081	0.0582	0.4189
	M3	**0.0406**	0.7324	0.0942	0.2089	0.5570
Beta	M1	**0.0062**	0.4618	0.1500	0.9775	0.7080
	M2	0.0529	0.6245	0.1485	0.4946	0.4215
	M3	0.1782	0.5437	0.1397	0.2575	0.5564

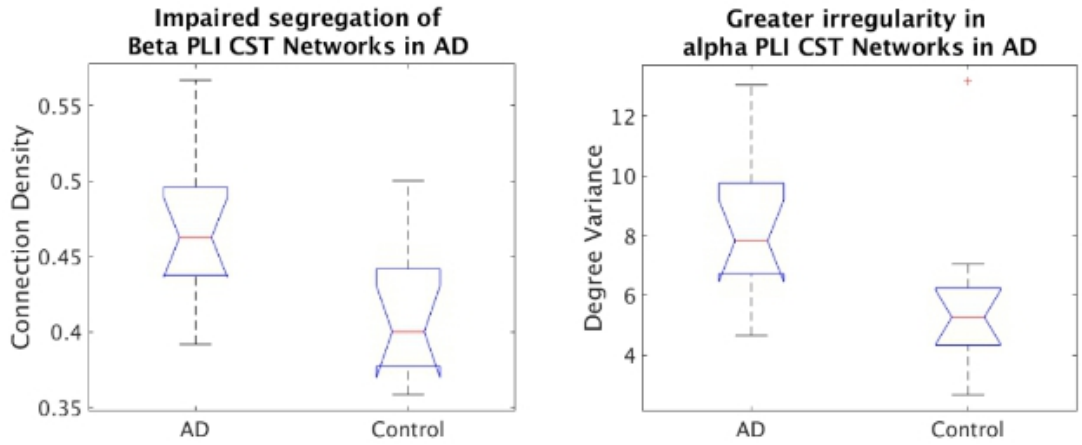

Figure 4.8: Box plots of connection density, left, and degree variance, right, for CST networks of AD and control in Beta and Alpha, respectively.

4.4.4 Density

The densities for the datasets used in this study are as in Table 4.7. For the USP we see both a dependency on network size, where the WCH networks are less dense with increasing size, and on the distribution of weights, where analysis of real datasets becomes implausible since connectivity, averaged over trials, creates a smaller spread and so redundant shortest paths. The MST and ECO are dependent on network size as is obvious from their formulations. The CST binarises the network consistently with a density between 0.3-0.5. From the models we notice that the higher the parameter, s, the less dense are the resulting CST networks. Network size appears to have much less effect, which provides evidence to suggest that the CST is dependent on topology, but not on network size.

4.5 Discussion

From the simulation results of complex hierarchy models we see from proportional thresholds that a larger density range is more effective than sparse models. This agrees with our hypothesis that complex hierarchical models contain a density of information beyond what sparse levels of binarisation can reveal. The fact that the real results for EEG datasets confirm the results in simulations provides further strength to the argument that EEG functional connectivity is highly hierarchically complex and so that sparsity is not always the best working hypothesis

Table 4.7: Mean and standard deviation (M ± SD) of densities of CST, USP, MST and ECO networks. WCH## = Weighted Complex Hierarchy model of size ##; s = strength parameter of WCH model; PLI## = PLI networks of size ##.

Dataset	Condition	CST	USP	MST	ECO
WCH16	$s = 0.1$	0.49 ± 0.04	0.53 ± 0.08	2/16	3/15
	$s = 0.15$	0.46 ± 0.05	0.59 ± 0.08		
	$s = 0.2$	0.44 ± 0.06	0.62 ± 0.08		
	$s = 0.25$	0.41 ± 0.06	0.64 ± 0.09		
	$s = 0.3$	0.40 ± 0.06	0.64 ± 0.05		
WCH32	$s = 0.1$	0.48 ± 0.02	0.45 ± 0.07	2/32	3/31
	$s = 0.15$	0.45 ± 0.03	0.50 ± 0.06		
	$s = 0.2$	0.42 ± 0.04	0.52 ± 0.07		
	$s = 0.25$	0.39 ± 0.04	0.53 ± 0.07		
	$s = 0.3$	0.37 ± 0.04	0.54 ± 0.08		
WCH64	$s = 0.1$	0.48 ± 0.01	0.38 ± 0.05	2/64	3/63
	$s = 0.15$	0.45 ± 0.03	0.42 ± 0.05		
	$s = 0.2$	0.41 ± 0.04	0.44 ± 0.05		
	$s = 0.25$	0.38 ± 0.04	0.45 ± 0.05		
	$s = 0.3$	0.36 ± 0.04	0.45 ± 0.06		
WCH128	$s = 0.1$	0.47 ± 0.01	0.33 ± 0.04	2/128	3/127
	$s = 0.15$	0.45 ± 0.02	0.35 ± 0.03		
	$s = 0.2$	0.40 ± 0.04	0.37 ± 0.03		
	$s = 0.25$	0.37 ± 0.04	0.37 ± 0.04		
	$s = 0.3$	0.35 ± 0.03	0.37 ± 0.06		
PLI16	Patient	0.47 ± 0.04	1 ± 0	2/16	3/15
	Control	0.41 ± 0.05	0.98 ± 0.05		
PLI30	Shape Left	0.41 ± 0.07	1 ± 0	2/30	3/29
	Shape Right	0.40 ± 0.06	1 ± 0		
	Bind Left	0.42 ± 0.06	1 ± 0		
	Bind Right	0.41 ± 0.08	1 ± 0		
PLI129	Eyes closed	0.35 ± 0.07	0.98 ± 0.04	2/129	3/128
	Eyes open	0.36 ± 0.07	0.98 ± 0.06		

for functional brain networks. Other evidence in EEG studies alluding to the benefit of medium density ranges has been documented [37, 49, 120]. For a physiological rationale for medium densities we can, for instance, regard function as not only emerging through physical connections as hypothesised in the sparsity hypothesis, but through globally interdependent synchronisations via rapid interregional communications. This interpretation holds in so far as the time necessary for such communication is available, as in the examples studied, and should be adjusted for cases in which only direct communications are feasible within the epoch analysed [130].

Binarised networks generally outperformed weighted approaches in our simulations. Furthermore, weighted network metrics should be used with caution. Particularly, we advise checking their correlations with the mean connection strength.

Other efforts looking to study the role of less strong connections in brain networks have also considered 'intermediate' thresholds by considering networks constructed not from strongest or weakest connections, but from connections within intermediate ranges of connectivity strength [118, 131]. However, studying such topologies is hindered by the fact that the true overlying hierarchical structure becomes hidden. Nodes having more edges in an intermediate level does not, for instance, indicate that that node is a hub, but rather that most of its connections lie within the given range. That is to say that intermediate connections maybe interesting to study, but constructing network topologies from them for analysis is rather obscure. The role of the weakest connections, or, if you will, topological gaps of brain connectivity is also an active area of research [131]. One can consider, in fact that medium density binarisation does much more to account for such features than sparse binarisations since these gaps become much more defined in higher densities.

As an important addition, the results show that the random topological attacks, rather than targeted topological attacks, are the most effective at deconstructing the topology of our simulations. This perhaps seems counter-intuitive, but can be explained by the fact that only the very top levels of the hierarchy are attacked in the targeted setting, whereas the topology in the remaining levels remains largely intact, maintaining the differences exhibited in the strength parameter, s, between the two sets of topology. In fact this agrees with previous functional connectivity studies which detailed the greater resilience of functional brain networks to targeted attacks [59, 132]. These simulations thus provide the clues as to how the hierarchical structure of functional brain networks play a vital role in this resilience.

The results for V in both the eyes open vs eyes closed and VSTM shape vs binding datasets combined can explain that more intensive stimulation (eyes open and binding) leads to a drop in network efficiency where more localised activity is required for higher functional processing [16, 41]. For the VSTM task, in particular, this drop in V for the more complex task may be explained by higher recruitment through a wider range of nodes. The results for the AD dataset both indicates the increased power in binarisation with the CST compared to other

approaches and highlights the importance of binarisation itself for distinguishing dysfunctional AD topology.

AD network studies, over varying platforms, network sizes and density ranges, have been found to show seemingly contrasting results found by Tijms *et al.* [17]. Particularly, they reported that these studies were at different density ranges, and in many cases the density range was simply not recorded. Importantly, no functional studies reported density ranges over 25%. Nonetheless, we note that our results are in agreement with a 149 node MEG PLI study by Stam *et al.* [32], showing lower clustering in AD than control (density not recorded). This is indicative of a move to a more random topology [52].

In terms of network size our simulations suggest that the larger the networks are, the more likely it is that topological differences will be picked up by commonly used metrics. This trend is bucked by the MST for which there is a marked drop off from 32 nodes to 128 nodes. This, however may be explained by the fact that at 32 nodes, the MST makes up $2/n = 6.25\%$ of all possible connections whereas at 64 nodes this percentage is 3.12% and for 128 nodes it is just 1.56% which is in line with the previous discussion that lower network densities can inhibit the ability to find topological differences.

The CST is presented here as a sensitive and powerful binarisation technique for network modelling of EEG phase-based functional connectivity. In simulations it performed to a high standard in all network sizes and topological comparisons as well as in robustness to topological attacks. This was echoed in the results of the real data sets where it was consistently able to identify differences in the presented conditions with not obviously correlated metrics. From the simulation results we can infer a large part of this ability to the density range in which the CST binarises the network. We must note, of course, that all of our real data were from EEG recordings and thus we are cautious of similar comparisons for e.g. fMRI. Indeed, we must acknowledge the limitations and narrow focus of this study for EEG PLI networks. Further, although we conjecture that hierarchical complexity of network topology may be behind the CSTs success, there remain unanswered questions and there is certainly scope for better topological thresholds to be developed based on such hypotheses. We hope this study will stimulate interesting discussions and inspire future research in this direction.

The MST is seen to be robust to fluctuations of the underlying network [44]. It holds appeal in studies where sparsity is desirable, showing utility in a number of other studies [33, 133–135], although it must be noted that in these studies it was not compared with other binarisation methods. In our study, however it appears ineffective and particularly so in larger networks which noticeably corresponds to the MST making up less and less of the connection density as the network grows. This concurs with a recent study where we argued that the robustness to fluctuations also means a poverty of information, supported by evidence from an EEG dataset of cognitive tasks [49].

The USP is the set union of those edges which form the shortest paths between all possible pairs of electrodes. Since, in general, all weights of a functional connectivity network lie between 0 and 1, it is likely that a large percentage of the shortest paths in the network will be constituted of just the single edge joining those nodes. Thus, we can expect very high density networks which only differ in topology by the weakest connections. Indeed, in the original paper [47] the authors did not implement any transformation of the weights and reported densities above 90%. To try and counter this unwanted outcome we used a negative exponential transform of the weights before extracting the union of shortest paths, however, in the end it appeared that this was limited in its ability to mitigate the flaws of this method. This was most apparent in the VSTM tasks where it turned out that every shortest path was just the edge between each pair of nodes, redundantly returning complete networks. We believe that further work would need to be done regarding the reliability of the USP in order to make it of use to the neuroscience community.

Although generally outperformed by the CST, we note an agreement with the introductory paper of the ECO threshold [50] that it generally outperforms the MST. Therefore, we would recommend it over the MST in cases where sparse densities appear more important. It was also able to detect differences in the left hemifield condition of the VSTM dataset. The fact that a mutually exclusive difference was found in the right hemifield condition with the CST suggests the interest in considering how different density ranges may reveal different topological traits of conditions. For example, one may conclude from these results that the backbone of the network is effected by binding in the left hemifield, but that in the right hemifield the binding effect is notable rather in the 'fleshed out' regions of the network.

As we have seen, it can be more beneficial to include a large number of edges in the network since it allows for richer information coming from the hierarchical relationships of lower degree nodes. This is important since it contradicts the implementation of statistical thresholds which looks for only connections whose strengths are statistically likely to be true connections. In such cases, it may be topologically more beneficial to keep more edges than what is allowed based on an arbitrary cut-off point of statistical significance. Other researchers look towards the integration of different density ranges, however such an approach will have a tendency towards diluting the potency of potential differences [111] or falling prey to the multiple comparison problem.

4.6 Conclusion

The hierarchical topologies of simulated weighted complex hierarchy models and several real datasets of EEG functional connectivity assessed from phase dependencies are found to be well characterised by non-arbitrary binarisations using the CST and arbitrary density binarisations in the range of 40%-50%. It is conjectured that this is due to their topologies in this range accounting for a wider range of hierarchical structure, i.e. not just the connectivity in the largest degree nodes. The CST and weighted networks were shown to be robust to random and targeted topological attacks when compared with MST, USP and ECO graphs. In three real datasets constituting varied neuroscientific questions, the medium density range which the CST occupies does indeed appear to be useful with other evidence showing that the ECO is useful in sparse densities. Considering both sparse and larger densities in tandem may prove a more effective way forward than either on their own. We were able to successfully identify different topologies in resting states, in VSTM cognitive tasks and in AD patients compared with control with a notable performance from the CST. This study also validates the WCH model as a sensitive tool for topological comparisons of great relevance to the EEG.

Chapter 5

Graph-Variate Signal Analysis: Framework and Applications to the EEG

The work in this chapter covers and extends the work published in Scientific Reports in February 2017 [4]. The extension provides a general framework within which the methods of this publication are framed, with which we explore and open up new possibilities for analysis. A pre-print is available online [3] and is currently under review.

5.1 Introduction

We wish to explore concretely the analysis of multivariate signals using graphs where each signal is assigned to a node of a graph. To do this we construct a unified framework encompassing both multivariate signals and graphs. With this, new modes of analysis become apparent. Notably, we can provide techniques which implement general bivariate functions on the signal which are then weighted by the graph topology, integrating the signal within a special graph adjacency matrix. This stems from studying the Dirichlet energy of a graph signal from GSP [78] which, in our new framework, can be noted as a particular case of our analysis. Exploring this connection, we show that matrix multiplications of an adjacency matrix with a signal vector is limited in the forms it can deliver and thus our framework provides a broader and more free setting for analysis of multivariate signals with graphs. We go on to describe interesting new ways to analyse multivariate signals using graphs including a new form of dynamic connectivity estimation and a form of network analysis conducted at the temporal resolution of the original signal. The former is particularly timely and promising in light of new efforts required for estimating dynamic connectivity from multivariate brain signals [69].

This theory provides a general approach for temporo-topological analysis of multivariate signals. Importantly, we provide a general framework for computing Graph-Variate Dynamic (GVD) connectivity and we provide reasonable solutions for frequency and phase-based connectivity in the form of correlation, coherence and PLI. We then demonstrate the power of our methodology by determining its ability to correctly identify the presence of correlations from various sizes of multivariate signals generated from an autoregressive process from which only a single couple of correlated signals exists. Furthermore, we demonstrate how the more refined analyses enabled by our generalisation provides greater accuracy in a simple randomly travelling spheroid detection problem than comparable approaches. Finally, we test the method on studying activated brain regions in terms of either amplitude or phase in neurophysiology in resting state and task-related activity.

5.2 Graph-variate signals

The framework is initialised with the introduction of a graph-variate signal which includes a multivariate signal associated with the node set similarly to how graph definitions usually include the weighted adjacency matrix associated with the edge set. Through this definition, one provides an object unifying multivariate signals and network science.

Definition 1. *We define the object* $\Gamma = (\mathcal{V}, \mathbf{X}, \mathcal{E}, \mathbf{W})$ *as a graph-variate signal where* $\mathcal{V}$ *is the set of vertices with* $|\mathcal{V}| = n$*;* $\mathbf{X} \in \mathbb{R}^{n \times p}$ *the multivariate signal indexed by* $\mathcal{V}$*;* $\mathcal{E} = \{(i,j) : i,j \in \mathcal{V}\}$ *the set of edges with* $|\mathcal{E}| = 2m$*; and* $\mathbf{W} = \{w_{ij}\}_{(i,j)\in\mathcal{E}} \in \mathbb{R}^{n \times n}$ *the weighted adjacency matrix encoding a relevant topology in which the multivariate signal is set. Then*

- $(\mathcal{V}, \mathbf{X})$ *is the node space composed of a matrix* $\mathbf{X}$ *whose first dimension is indexed by the node set* $\mathcal{V}$ *and second dimension is indexed by a sequential characteristic of activity at the nodes, typically time.*
- $(\mathcal{E}, \mathbf{W})$ *is the edge space composed of a weighted matrix* $\mathbf{W}$ *indexed by the edge set* $\mathcal{E}$*.*
- Γ *constitutes the graph space of the combined node and edge spaces where vertices and edges joining those vertices are determined by the node labels* $\{1, \ldots, n\}$*.*

Note that, in the simplest case, $\mathbf{W}$ is fixed, but it is straightforward to consider an extension in which $\mathbf{W}$ also changes over time. Nevertheless, in the case where $\mathbf{W}$ is estimated from

the signal itself, as we shall later study, it is recommended that $\mathbf{W}$ represent the longest time window possible to increase robustness. The node space, being that which contains the activity at the nodes, frames the standard analysis of multivariate signals. Indeed, this is formalised by a general node function, $F_{\mathcal{V}}$, defined on the node space $(\mathcal{V}, \mathbf{X})$ as

$$\begin{array}{rccc} F_{\mathcal{V}}: & \mathbb{R}^{n\times p} & \rightarrow & \mathbb{R}^{m\times q} \\ & \mathbf{X} & \mapsto & F_{\mathcal{V}}(\mathbf{X}), \end{array} \tag{5.1}$$

and, when applied in a single instance to pairs of channels at sample s, as

$$\begin{array}{rccc} F_{\mathcal{V}}: & \mathbb{R}^{1\times 2} & \rightarrow & \mathbb{R} \\ & [x_i(s), x_j(s)] & \mapsto & F_{\mathcal{V}}(x_i(s), x_j(s)). \end{array} \tag{5.2}$$

Useful examples of such functions where $n = m$ and $p = q$ include weight thresholds and a spectral filtering function, e.g. for bandpassing the signal in a frequency band of interest.

The edge space, on the other hand, is a topological space whose elements are the unlabelled isomorphism classes of graphs of size n. This is where one finds the standard analysis of networks. A function $F_{\mathcal{E}}$ on the edge space $(\mathcal{E}, \mathbf{W})$ is defined on $\mathbb{R}^{n\times n}$ as

$$\begin{array}{rccc} F_{\mathcal{E}}: & \mathbb{R}^{n\times n} & \rightarrow & \mathbb{R}^{m\times l} \\ & \mathbf{W} & \mapsto & F_{\mathcal{E}}(\mathbf{W}). \end{array} \tag{5.3}$$

Such functions can be thresholds when $n = m = l$, global network indices, such as transitivity or characteristic path length,when $m = l = 1$ and local network indices, such as the local clustering coefficient or betweenness centrality, when $m = n$ and $l = 1$. These are necessarily all invariants under graph isomorphisms- individuality of nodes is not considered.

The following definitions will be useful for specificity in the rest of this section.

Definition 2. *An edge dimension preserving function, $\bar{F}_{\mathcal{E}}$, maps the adjacency matrix, $\mathbf{W} \in \mathbb{R}^{n\times n}$, to a new matrix $\tilde{\mathbf{W}} \in \mathbb{R}^{n\times n}$.*

Definition 3. *A node dimension preserving function, $\bar{F}_{\mathcal{V}}$, maps the multivariate signal, $\mathbf{X} \in \mathbb{R}^{n\times p}$, to a new signal $\tilde{\mathbf{X}} \in \mathbb{R}^{n\times p}$.*

This framework and GSP both share in common a difference from temporal networks– where events occur at specific times between nodes (i.e. at edges)– in that the graph edges are fixed and we are instead concerned with the study of signal activity occurring at the nodes. A key distinction to make here is that in this framework the signals are sampled with respect to a sequential characteristic, such as time, but the interdependencies are known or estimated and encoded in a graph topology, whereas in GSP the signal is sampled and thus processed over a graph topology. By theoretical considerations we will demonstrate that the former is more flexible for analysing multivariate signals.

We shall now consider how node and edge spaces can be combined to produce meaningful analyses for the graph-based analysis of multivariate signals. Important considerations of such operations pertain to how edge space operations can act on the node space and, reciprocally, how node space operations can act on the edge space. In the usual graph sense it is required that these operations preserve the inner dimensions whose size is the same as the node set, n, before acting on their reciprocal space.

5.2.1 Edge-dependent operations acting on the node space

Since the inner-dimensions of the edge space and node space agree, the output of any edge-dimension preserving function together with the usual matrix multiplication, $\cdot$, provide useful operations which act on the node space, $(\mathcal{V}, \hat{\mathbf{X}})$:

$$\begin{array}{rccc} \bar{F}_{\mathcal{E}}(\mathbf{W})\cdot: & \mathbb{R}^{n\times p} & \rightarrow & \mathbb{R}^{n\times p} \\ & \mathbf{X} & \mapsto & \bar{F}_{\mathcal{E}}(\mathbf{W})\cdot\mathbf{X}. \end{array} \tag{5.4}$$

We thus realise that $\bar{F}_{\mathcal{E}}(\mathbf{W})\cdot$ is in fact a node dimension preserving function. Some of the simplest examples include the weighted adjacency matrix, $\mathbf{W}$, and the graph Laplacian, $\mathbf{L}$. Indeed, this property is exploited to formulate the various aspects of GSP where important definitions involve pre-matrix multiplication of the graph signal by matrices derived from graphs. For example, the GFT treats the eigenvectors of the Laplacian or the graph adjacency matrix as a basis for the decomposition of graph signals into graph frequency components. The lth eigenvector produces the lth frequency component of the graph signal, $\mathbf{x} \in \mathbb{R}^{n\times 1}$, defined as $\mathbf{u}_l \cdot \mathbf{x}$. Similarly, graph convolution, translation, modulation and graph wavelets

can be formulated as matrix multiplication on linear components of the graph signal. Further, polynomials of the adjacency and Laplacian matrices are implemented to construct graph signal filters in GSP in [79] and [78], respectively, which are then matrix multiplied by the graph signal. We will now evolve a new analysis of graph-variate signals by reciprocally considering node space functions acting on adjacency matrices.

5.2.2 Node-dependent operations acting on the edge space

Because the edge space is composed of pairs of elements in the node space, when combining the output of node space functions with the adjacency matrix it is most sensible to impose that the elements acting on the weight w_{ij} be bivariate functions of the signal at nodes i and j. In this manner, we define graph-variate signal analysis.

Definition 4. *Graph-variate signal analysis is an all-to-all bivariate analysis of the signal* $\mathbf{X}$ *weighted by the corresponding edge weights in* $\mathbf{W}$.

Graph-variate signal analysis is facilitated by formulating a tensor, $\underline{\mathbf{J}} \in \mathbb{R}^{n\times n\times p}$, which is the output of a node space function defined as

$$J_{ijt} = \begin{cases} F_{\mathcal{V}}(x_i(t), x_j(t)), & i \neq j \\ 0, & i = j, \end{cases} \tag{5.5}$$

for some node space function $F_{\mathcal{V}}$.

In order to fully encode the graph-variability we consider both the edge and node spaces,

$$(\mathbf{W} \circ \underline{\mathbf{J}}_{(t)})_{ij} = \begin{cases} w_{ij} F_{\mathcal{V}}(x_i(t), x_j(t)), & i \neq j \\ 0, & i = j, \end{cases} \tag{5.6}$$

where $\underline{\mathbf{J}}_{(t)}$ denotes the tth $n \times n$ matrix of $\underline{\mathbf{J}}$ and $\circ$ is the mode-k Hadamard product. This way, w_{ij}, which encodes the relationship between nodes i and j, is multiplied by the relevant node space function on x_i and x_j.

It is also useful to define a new operator which allows node space operations to act on the edge space to provide local graph-variate analysis for each node.

Definition 5. *For a matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ and 3D tensor $\underline{\mathbf{B}} \in \mathbb{R}^{n\times n\times p}$, composed of the p $n \times n$ matrices $\{\underline{\mathbf{B}}_{(k)}\}_{t=1}^{p}$, their **signal product**, $\mathbf{A} \diamond \underline{\mathbf{B}}$, is a matrix whose kth column is the vector $[\sum_j A_{ij}B_{jik}]_{i=1}^{n}$, which is the dot product of the ith rows of $\mathbf{A}$ and the ith columns of $\underline{\mathbf{B}}_{(k)}$.*

Then

$$(\mathbf{W} \diamond \underline{\mathbf{J}})_{it} = \sum_{j=1}^{n} w_{ij} F_{\mathcal{V}}(x_i(t), x_j(t)). \tag{5.7}$$

A special case of this is GSP's node gradient formula [78] where $F_{\mathcal{V}}(x_i(t), x_j(t)) = (x_i(t) - x_j(t))^2$.

This operator has the interesting property of providing a reciprocal approach (up to linear combinations) for the matrix multiplication operator which allows edge space operations to act on the node space. From this, the limitations of the GSP framework for our ends are made most apparent. It is straightforward to note that node space functions $x_j(t)$ and $x_i(t)-x_j(t)$ are solutions for $F_{\mathcal{V}}$ in (5.7) to the equations $\mathbf{W}\diamond\underline{\mathbf{J}} = \mathbf{W}\cdot\mathbf{X}$ and $\mathbf{W}\diamond\underline{\mathbf{J}} = \mathbf{L}\cdot\mathbf{X}$, respectively. The limitations of adjacency matrix multiplication with signal vectors are revealed in the following.

Proposition 1. *For the output of an edge dimension preserving function $\bar{F}_{\mathcal{E}}(\mathbf{W})$ and of a node dimension preserving function $\bar{F}_{\mathcal{V}}(\mathbf{X})$,*

$$\bar{F}_{\mathcal{E}}(\mathbf{W}) \cdot \mathbf{X} = \mathbf{W} \diamond \bar{F}_{\mathcal{V}}(\mathbf{X}) \tag{5.8}$$

if and only if $\bar{F}_{\mathcal{V}}(\mathbf{X}) = a_{ij}x_i(t) + a_{ji}x_j(t)$ for some constants $a_{ij}, a_{ji} \in \mathbb{R}$, and

$$\bar{F}_{\mathcal{E}}(\mathbf{W}) = \begin{bmatrix} \sum_j a_{1j}w_{1j} & a_{21}w_{12} & \dots & a_{n1}w_{1n} \\ a_{12}w_{21} & \sum_j a_{2j}w_{2j} & \dots & a_{n2}w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n}w_{n1} & a_{2n}w_{n2} & \dots & \sum_j a_{nj}w_{nj} \end{bmatrix}. \tag{5.9}$$

Proof. We first note that matrix multiplication with $\mathbf{X}$ is linear on the entries of $\mathbf{X}$ thus we cannot consider equating $\bar{F}_{\mathcal{E}}(\mathbf{W}) \cdot \mathbf{X}$ to a graph weighted non-linear node space function– one cannot obtain elements $x_i(t)^p$ for $p > 1$. Further, since each element of $\bar{F}_{\mathcal{E}}(\mathbf{W})$ is multiplied by an element of $\mathbf{X}$ and each element of $\bar{F}_{\mathcal{V}}(\mathbf{X})$ is multiplied by an entry of $\mathbf{W}$, there can be no constants in either function.

Now, in the linear case without constants for $\mathbf{x} \in \mathbb{R}^{n\times 1}$,

$$(\tilde{\mathbf{W}} \cdot \mathbf{X})_{ti} = \sum_{j=1}^{n} w_{ij}(a_{ij}x_i(t) + a_{ji}x_j(t))$$

$$\iff \tilde{w}_{ij} = \begin{cases} \sum_{p=1}^{n} a_{ip}w_{ip} & i = j \\ a_{ji}w_{ij} & i \neq j, \end{cases} \tag{5.10}$$

for coefficients $a_{ij} \in \mathbb{R}$, satisfying the proposition. □

This goes beyond realising that $\mathbf{W}$ and $\mathbf{L}$ as operators for graph signals obtain only special forms of GVD connectivity, but further that no graph matrix can be constructed to allow for non-linear node functions via matrix multiplication with graph signal vectors.

5.2.3 Graph-variate networks

Interestingly, from (5.6) we note that $\underline{\mathbf{\Delta}}_{(t)} = \mathbf{W} \circ \underline{\mathbf{J}}_{(t)}$ itself takes on a weighted adjacency matrix form and thus the tensor $\underline{\mathbf{\Delta}} \in \mathbb{R}^{n\times n\times p}$ is a multi-layer network of sequentially related graphs. This is useful as we can then explore topological characteristics of a graph-variate signal at every sample. In classical network science, there are many methods proposed to analyse the topology of a graph by applying operations in the edge space, that is, on the edge weight matrix, $\mathbf{W}$. Such methods provide important insights and classifications of the interdependent relationships of the underlying objects [6]. In our experiments, we will implement a simple example of a local clustering coefficient, C_{loc}, of node i at time t, defined for the graph-variate signal as

$$C_{loc}(i,t) = \sum_{j,k=1}^{n} \Delta_{ijt}\Delta_{ikt}\Delta_{jkt} = (\underline{\mathbf{\Delta}}^3_{(t)})_{ii}. \tag{5.11}$$

Implementing network science on $\underline{\mathbf{\Delta}}$ could, for example, provide insights into rapid fluctuations in the topological relationships of the signals or, as we will focus on in Section 5.4.2, could be used to develop techniques based jointly on pairwise signal dependencies and their spatial distances.

5.3 Graph-variate dynamic connectivity analysis

Now we turn out attention to the special case in which the graph weights encode pairwise dependencies which have been estimated using the whole multivariate signal itself. Specifically, for important connectivity measures, we demonstrate how appropriate consideration of the node space function in graph-variate signal analysis provides a sample resolution analysis of dynamic connectivity. The following makes use of the instantaneous amplitude and phase components of the analytic representation of the univariate signals $\mathbf{x}_i$, of $\mathbf{X}$, $x_i^a(t) = s_i^a(t)e^{j\phi_i}(t)$.

We define a function of connectivity of pairwise channels, $\{\mathbf{x}_i, \mathbf{x}_j\}$, in $\mathbf{X}$ as

$$\begin{array}{rccc} H_{\mathcal{V}}: & \mathbb{R}^{2\times p} & \rightarrow & \mathbb{R} \\ & \{\mathbf{x}_i, \mathbf{x}_j\} & \mapsto & H_{\mathcal{V}}(\mathbf{x}_i, \mathbf{x}_j), \end{array} \tag{5.12}$$

where $H_{\mathcal{V}}$ is a node function of the entire signals $\mathbf{x}_i$ and $\mathbf{x}_j$ and $H_{\mathcal{V}}(\mathbf{X}) = \mathbf{C} \in \mathbb{R}^{n\times n}$ such that

$$c_{ij} = \begin{cases} H_{\mathcal{V}}(\mathbf{x}_i, \mathbf{x}_j), & i \neq j \\ 0, & i = j. \end{cases} \tag{5.13}$$

If the function is symmetric, then the matrix $\mathbf{C}$ is regarded as the weighted adjacency matrix of an undirected graph. Otherwise the graph is directed. We focus only on the undirected case here, however directed graphs may also be considered. Note that $\mathbf{C}$ is equivalent to $\mathbf{W}$ in the general case.

We define GVD connectivity as a graph-variate signal analysis with a suitably chosen node space function (5.5) weighted by a connectivity adjacency matrix, $\mathbf{C}$, derived from the signal. GVD connectivity then takes the form

$$\theta(\mathbf{x}_i, \mathbf{x}_j, t) = \begin{cases} c_{ij}F_{\mathcal{V}}(x_i(t), x_j(t)), & i \neq j \\ 0, & i = j. \end{cases} \tag{5.14}$$

By an appropriate choice of $F_{\mathcal{V}}$ this can reveal information of the significance of specific

points in time of the connectivity over which the adjacency matrix is constructed. From another perspective, the connectivity adjacency matrix acts as a filter for extracting useful information from the node space function, $F_{\mathcal{V}}$- strong connectivity implies those vertices are sharing or communicating important information consistently, thus these connections amplify the function at those vertices, whereas weak connectivity implies the opposite and suppresses the function at those vertices [4]. The outline of GVD connectivity is very flexible, but this also means that in its application it requires a carefully considered formulation, as we undertake here, to avoid data-dredging.

A particularly useful analysis for exploring the GVD connectivity associated with a particular node is the *node GVD connectivity*:

$$\theta_i(\mathbf{X}, t) = \sum_{j=1}^{n} c_{ij} F_{\mathcal{V}}(x_i(t), x_j(t)). \tag{5.15}$$

We will use this a number of times in our experiments. The operator which extracts the vector of node temporal connectivities is defined in (5.7).

We can also easily carry out analysis of modules, *module GVD connectivity*:

$$\theta_{\mathcal{V}_a}(\mathbf{X}, t) = \sum_{i \in \mathcal{V}_a} \sum_{j=1}^{n} c_{ij} F_{\mathcal{V}}(x_i(t), x_j(t)). \tag{5.16}$$

following the disjoint composition of the underlying node set into q modules,

$$\mathcal{V} = \bigsqcup_{i=1}^{q} \mathcal{V}_i. \tag{5.17}$$

Further, we can analyse activity between modules

$$\theta_{\mathcal{V}_a, \mathcal{V}_b}(\mathbf{X}, t) = \sum_{i \in \mathcal{V}_a} \sum_{j \in \mathcal{V}_b} c_{ij} F_{\mathcal{V}}(x_i(t), x_j(t)). \tag{5.18}$$

for two disjoint modules $\mathcal{V}_a$ and $\mathcal{V}_b$. Modular analysis is particularly useful for understanding activity related to brain regions.

Here we present node functions for three pertinent examples of connectivity adjacency matrices- correlation, coherence and PLI- of which the reader may refer back to Section 2.3

for further details. For clarity of exposition, in each case we will first present the formulae for these connectivity estimates before going on to describe the chosen node space functions to compute GVD connectivity.

5.3.1 Correlation

Taking the connectivity estimate as the correlation coefficient we have

$$c_{ij} = \frac{\sum_{t\in T}(x_i(t) - \bar{\mathbf{x}}_i)(x_j(t) - \bar{\mathbf{x}}_j)}{\sqrt{\sum_{t\in T}(x_i(t) - \bar{\mathbf{x}}_i)^2}\sqrt{\sum_{t\in T}(x_j(t) - \bar{\mathbf{x}}_j)^2}} \tag{5.19}$$

where T is the epoch of interest, and $\bar{\mathbf{x}}_i$ is the mean of the values over time of the node i. As a preliminary formulation, we will consider as node function:

$$\theta(\mathbf{x}_i, \mathbf{x}_j, t) = c_{ij}(\tilde{x}_i(t) - \tilde{x}_j(t))^2, \tag{5.20}$$

derived from the Dirichlet energy form from GSP [78], where $\tilde{\mathbf{x}}_i(t)$ is the normalised signal over the node space, i.e.

$$\tilde{x}_i(t) = \frac{x_i(t) - \bar{\mathbf{x}}(t)}{\sqrt{\frac{1}{n-1}\sum_{k=1}^{n}(x_k(t) - \bar{\mathbf{x}}(t))^2}}, \tag{5.21}$$

where $\bar{\mathbf{x}}(t) = \frac{1}{n}\sum_{k=1}^{n} x_k(t)$ is the mean over vertices of the signal at time t. Notably, the entries of the matrix may be negative which is an important principle for maintaining the anti-correlative information, as we shall see.

With some careful thought we see that even this simple form can be regarded as an appropriate function for correlation. Consider the relation

$$c_{ij}(\tilde{x}_i(t) - \tilde{x}_j(t))^2 \tag{5.22}$$

where c_{ij} is the correlation, possibly negative, between signals i and j over epoch T. When correlation between signals is low, the magnitude of (5.22) is small and we cannot infer much from the signals. Now we consider when the magnitude of the correlation is large. There are four cases to consider: i) $c_{ij} > 0$ and $(\tilde{x}_i(t) - \tilde{x}_j(t))^2$ is large; ii) $c_{ij} > 0$ and $(\tilde{x}_i(t) - \tilde{x}_j(t))^2$ is small; iii) $c_{ij} < 0$ and $(\tilde{x}_i(t) - \tilde{x}_j(t))^2$ is large; iv) $c_{ij} < 0$ and $(\tilde{x}_i(t) - \tilde{x}_j(t))^2$ is small.

The first case says there is a strong positive correlation, c_{ij}, between signals i and j and x_i and x_j at time point t are dissimilar. Since high positive correlation denotes that the signal amplitudes behave similarly, this large positive value indicates a likely discrepancy between the connectivity information over the epoch T and the behaviour of the signals at time point t. Here the output is positive and large. The second case, on the other hand says that there is a strong positive correlation between the signals i and j and the signals are similar at t. This, in contrast to the first case, gives two agreeing components and the output is positive and small. The third case says there is a strong negative correlation, c_{ij}, between signals i and j and there is a large difference between signals i and j at time t. Since high negative correlation denotes that the signal amplitudes behave dissimilarly, this large positive value indicates agreement between the connectivity information over the epoch T and the behaviour of the signals at time point t. The output is negative and large. The fourth case, on the other hand, says that there is a strong negative correlation between the signals i and j and the signals are similar at t. This, in contrast to the third case, gives two disagreeing components and the output is negative and small.

Consider summing a variety of these components, for instance over part or whole of the graph. If the result is large and positive, we can say that the activity of the signals at time t is mostly in disagreement with the connectivity over T. On the other hand, if the output is large and negative, we can say that the activity of the signals at time t is mostly in agreement with the connectivity over T.

On the other hand, from our generalised setting, we can consider an instantaneous correlation deriving more directly from (5.19):

$$\theta(\mathbf{x}_i, \mathbf{x}_j, t) = c_{ij}|(x_i(t) - \bar{\mathbf{x}}_i)(x_j(t) - \bar{\mathbf{x}}_j)|, \tag{5.23}$$

where the node space function here can be understood as a measure of instantaneous correlation coefficient at the time point t. These methods shall be compared in simulations and real data to help reveal the benefits of using more nuanced functions than the classical Dirichlet form.

5.3.2 Coherence

The coherence of two channels is a function of frequency, ω, and can be interpreted as a correlation of signal component at ω of the channels. For a chosen frequency band we have

$$c_{ij} = \sum_{\omega \in \Omega} \frac{|P_{\mathbf{x}_i \mathbf{x}_j}(\omega)|^2}{P_{\mathbf{x}_i \mathbf{x}_i}(\omega) P_{\mathbf{x}_j \mathbf{x}_j}(\omega)}, \tag{5.24}$$

as in Section 2.3.

Similarly as for correlation, we will consider both the squared difference and instantaneous correlations of the instantaneous amplitudes to compute GVD connectivity for coherence, after bandpassing in the frequency range of interest, i.e.

$$\theta(\mathbf{x}_i, \mathbf{x}_j, t) = c_{ij}(s_i^a(t) - s_j^a(t))^2. \tag{5.25}$$

and

$$\theta(\mathbf{x}_i, \mathbf{x}_j, t) = c_{ij}|(s_i^a(t) - \bar{\mathbf{s}}_i^a)(s_j^a(t) - \bar{\mathbf{s}}_j^a)|, \tag{5.26}$$

respectively. Coherence cannot be negative, thus it is a more straightforward case than correlation- high coherence and large differences in the instantaneous amplitudes can be taken generally as a contrast of information; whereas a small difference in amplitudes implies agreement of information. Thus large GVD connectivity implies some notable epoch of interest in the given time window with the underlying long-term connectivity.

5.3.3 Phase-lag index

The PLI [31] measures the consistent phase differences between time-series, indicating lead/lag dependencies. As a connectivity estimate, we write

$$c_{ij} = |\langle \text{sgn}(\phi_i(t) - \phi_j(t)) \rangle|. \tag{5.27}$$

We choose $F_{\mathcal{V}}$ for phase-based connectivity indexes as the sign of the phase difference of the

signals stemming directly from (5.27), giving

$$\theta(\mathbf{x}_i, \mathbf{x}_j, t) = c_{ij}\text{sgn}(\phi_i(t) - \phi_j(t)). \tag{5.28}$$

Because of the negative symmetry of this function, the global GVD connectivity of the system at time t is

$$\begin{aligned}\sum_{i,j} \theta(\mathbf{x}_i, \mathbf{x}_j, t) &= \sum_{i<j} (\theta(\mathbf{x}_i, \mathbf{x}_j, t) + \theta(\mathbf{x}_j, \mathbf{x}_i, t)) \\ &= \sum_{i<j} (\theta(\mathbf{x}_i, \mathbf{x}_j, t) - \theta(\mathbf{x}_i, \mathbf{x}_j, t)) = 0.\end{aligned} \tag{5.29}$$

However, summing over a subset of these elements, for example, only over those edges relating to a given node or subset of nodes, would reveal the strength and general nature of the node(s) to lead (positive) or lag (negative) in the network at the given epoch. In experiments we will apply these GVD connectivity functions to several simulated and real datasets to provide document of their usefulness.

5.4 Experiments

We now demonstrate these methods in simulations and real data sets. An autoregressive model is implemented first to illustrate the broad idea and benefit of graph-variate signal analysis before we explore the ability of GVD connectivity to correctly discover differences between two large datasets which differ only by the presence (and lack thereof) of a single correlated couple (Section 5.4.1). To test the effectiveness of temporal network clustering coefficient metric (5.11), we devise a simple regime to detect a spheroid travelling over a 3D grid (Section 5.4.2). We then apply our techniques to real high complexity datasets of EEG brain functional connectivity to provide evidence of the benefits delivered by a graph-variate analysis approach (Sections 5.4.3-5.4.5).

5.4.1 Detecting correlated sources

We generate 5 realisations, 1×1000 vectors $\{\mathbf{z}_i\}_{i=1}^{5}$, of a stationary autoregressive process with governing equation

$$z(t) = 0.5 + 0.7z(t-1) + 0.25z(t-2) + \epsilon, \tag{5.30}$$

where $\epsilon \sim \mathcal{N}(0, 0.1)$ and consider the multivariate signal

$$\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \\ \mathbf{z}_3 \\ \mathbf{z}_4 \\ \mathbf{z}_5 \end{bmatrix}, \tag{5.31}$$

so that all $\mathbf{x}_i$ are the average of two realisations of (5.30); $\mathbf{x}_1$ and $\mathbf{x}_2$ are correlated via the information in $\mathbf{z}_1$; and $\mathbf{x}_3$ is independent of $\mathbf{x}_1$ and $\mathbf{x}_2$. Fig 5.1 shows the computation of instantaneous correlation coefficients and corresponding node GVD connectivity computed using correlation coefficient (5.23) over the entire signal. The corresponding graph weights are $w_{12} = 0.6934, w_{13} = -0.0576, w_{23} = 0.0943$. Node GVD connectivity (bottom) is computed over 5 samples in non-overlapping windows. The corresponding short-term graph weights computed over 5 samples and the un-weighted instantaneous correlation are shown in the 2nd and 3rd plots, respectively. Unsubstantiated dependencies are produced using the short-term graph weight and instantaneous correlation methods where often the three components are roughly equivalent. GVD connectivity, on the other hand suppresses the uncorrelated data using the long-term connectivity estimates and the prevailing information comes forth from the truly correlated data relating to edge $(1, 2)$. This is most obviously seen in comparing instantaneous correlation (third) with GVD connectivity (bottom), where the signals are identical except that GVD connectivity weights them by long-term correlations, hence the yellow, $(2, 3)$, and orange, $(1, 3)$, time-series are suppressed relative to the blue time-series, $(1, 2)$.

Note that the amount of data required to measure statistical moments of a random variable

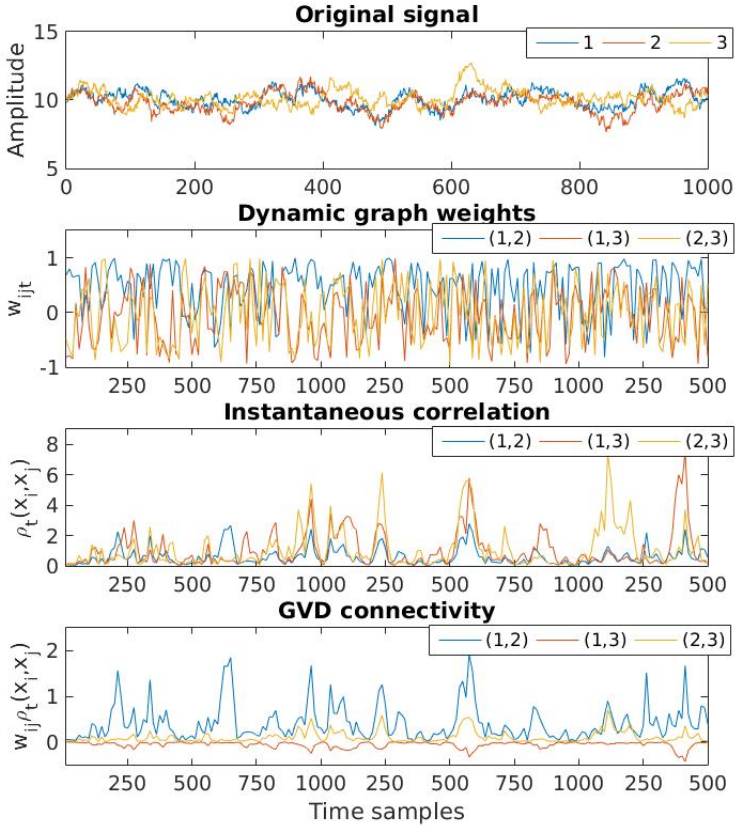

Figure 5.1: The original signal (top), dynamic graph weights (second) Instantaneous correlations (third) and corresponding GVD connectivity (bottom) of edges as shown in the legend. The benefit of long-term graph weights is evident, where the GVD connectivity correctly emphasises important information (that related to edge (1,2)).

that is correlated in time to a given accuracy is dependent on the amount of temporal correlation present. What theoretical results there are exist for Gaussian random signals that are independent from sample to sample in time. If we use N samples of such to estimate the mean of a Gaussian variable then the estimate is also a random variable with the same mean (i.e. unbiased) but the variance of the estimate is $1/N$ the variance of the original random signal. When the signal is correlated in time even more data is required to estimate the mean to a given accuracy. This provides the statistical argument for what is happening in this toy example and the reason for using long-term dependency estimates to ameliorate attempts at measuring transient dynamic dependencies.

We now extend this to quantitatively assess the ability to determine a single couple of correlated signals from increasingly large sets of signals. Following the same autoregressive process as (5.30), we generate $2 \times h$ realisations for $h = 2, 4, 8, 16, \ldots, 512$. Then two sets of signals

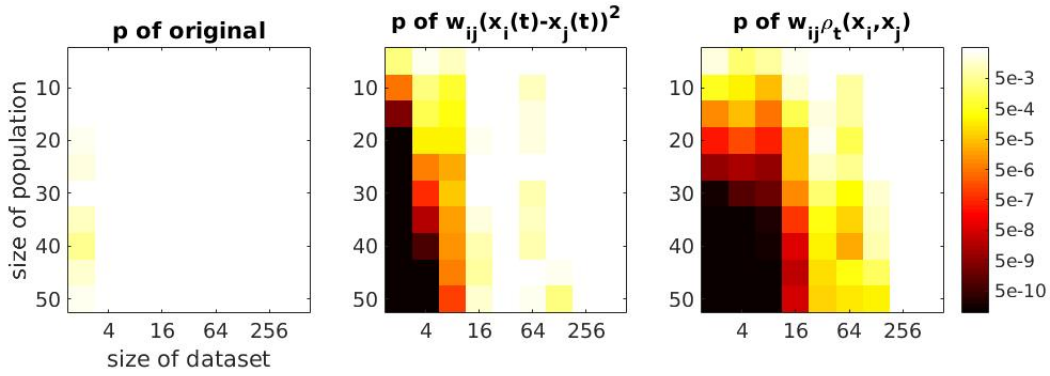

Figure 5.2: The p-values of one-sample t-tests of correctly identified correlated sources for different sizes of dataset (x-axis) and population (y-axis) for the original signal (left) and GVD node connectivity with squared difference (middle) and instantaneous correlation (right). White indicates a non-significant difference, black indicates a p-value value smaller than 5×10^{-10}.

are formed. The first uncorrelated set takes the average of each consecutive disjoint couple of realisations as the multivariate signal $\mathbf{X} \in \mathbb{R}^{h \times 1000}$. The second set is almost the same except the second signal is formed from the 1st and 4th (rather than 3rd and 4th) realisations so as to be correlated with the first signal. We generate populations of such multivariate signals of sizes $5, 10, 15, \ldots, 50$ to track effects due to population size. We then compute the difference between the uncorrelated and correlated original signals and GVD connectivity analysis using (5.20) and (5.23) and sum over time. We implement a one-sample t-test on the null hypothesis that the population values have a zero mean with significance indicating rejection of the null hypothesis at the $\alpha = 0.05$ level. The results for each population and signal size are shown in Fig 5.2.

The values for the original signals are provided for reference since we do not expect them to perform well given that they rely only on magnitudes. The results clearly indicate that GVD connectivity using instantaneous correlation has greater sensitivity to differences than using squared difference. Specifically, we can state that GVD connectivity with instantaneous correlation can correctly and reliably identify differences in the autoregressive processes with a population size of 25 or greater with at least 128 signals. Squared difference can detect differences in 128 signals only with a population size of 50. However, even with a population of 50 this is not reliably since it fails to detect the difference in the 32 signal case.

5.4.2 Spheroid travelling randomly on a 3D grid

We construct a $10\times10\times10$ grid in Euclidean space where each point corresponds to a univariate signal. A weighted connectivity graph is formed from the inverse distance, computed as $w_{ij} = \exp(-d_{ij}^2/4)$, between the intersecting points in the grid and amplitudes are distributed normally at random to the vertices as $\mathcal{N}(0, 0.3)$. At time t, the amplitude at node i is increased arbitrarily by an amount, δ, and amplitudes at those vertices one unit away from i are increased by $\frac{3}{4}\delta$. At time $t+1$, again amplitudes are assigned normally at random but one of the vertices assigned $\frac{3}{4}\delta$ increased amplitude at time t is now randomly selected for a δ amplitude increase and its closest neighbours are now assigned $\frac{3}{4}\delta$. We can liken this to a spheroid travelling randomly across a grainy image. This process continues for 1000 time steps for values of δ ranging from 0.1 in steps of 0.1 up to 0.9. We now consider the appropriate node space function to use in this scenario. The randomness of movement means that using approaches which try to assess a direction, such as Kalman filtering, are of little value. Thus, a more basic maximisation approach is adopted. Considering that higher amplitudes close together should produce high values, we implement graph-variate signal analyses using a multi-layer graph, $\underline{\mathbf{\Delta}}$, with a node space function which takes the average of each signal pair so that:

$$\Delta_{ijt} = \tfrac{1}{2} w_{ij}(x_i(t) + x_j(t)). \tag{5.32}$$

We then calculate the weighted clustering coefficient, C_{loc}, from (5.11), at each node at each point in time. The task is then to detect the spheroid at each point in time. We compare with simply choosing the node with highest amplitude and also by implementing graph filtering approaches based on the graph adjacency matrix with self-loops, $\hat{\mathbf{W}} = \mathbf{I} + \mathbf{W}$ [79], and the graph Laplacian [78], as well as using the heat kernel, $e^{-\tau\mathbf{L}}$ [78]. That is, at time t, we select the highest value of the vectors $\hat{\mathbf{W}}\mathbf{X}(t)$, $\mathbf{L}\mathbf{X}(t)$, and $e^{-\mathbf{L}}\mathbf{X}(t)$ and also the cubed versions $\hat{\mathbf{W}}^3\mathbf{X}(t)$, $\mathbf{L}^3\mathbf{X}(t)$, and $e^{-3\mathbf{L}}\mathbf{X}(t)$ to compare a simple GVD connectivity approach with some standard GSP approaches.

We take the largest clustering coefficients as the measure to detect the spheroid at each point in time and compare with just taking the highest amplitude value of the signal and the highest value of the outlined GSP approaches. Fig 5.3 details the number of correctly identified

spheroid centres (left) and the number of identifications at any point of the spheroid (right), i.e. within one grid square of centre, for each δ. Our approach using C_{loc} (green) achieves best results in 7/9 cases in the former and in all cases in the latter. Our approach shows best overall results in both, see Table 5.1, being one percentage point clear of the next best in detecting the centre and nearly ten percentage points clear of the next best in detecting any part of the spheroid. Of the GSP approaches, the best are the single adjacency matrix filter $\hat{\mathbf{W}}$ (Fig 5.3, dark blue) and the heat kernel $e^{-3\mathbf{L}}$ (orange), which perform relatively well in detecting the centre point. However, they fair much less well when taking into account the sides of the spheroid, where they do not fair much better than the default maximum amplitude approach (black). Since $\mathbf{W}$ and $\mathbf{L}$ fair better than their cubic versions, we know that the improvement noted by the clustering coefficient method is not down to the cube of the graph distance information resulting from (5.11). Indeed, the graph topology being emphasised in higher powers of the adjacency matrix and graph Laplacian leads to a decrease in the relevant information contained in the signal, which is not dependent on the topology of the grid.

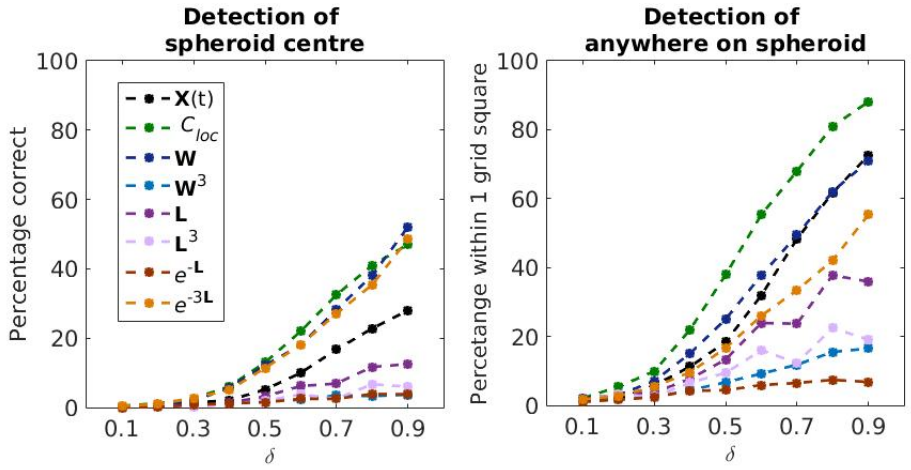

Figure 5.3: Total of correct guesses, left, and guesses anywhere on the spheroid, right, out of 1000 time points using amplitude height only (max. amp.), signal function graph clustering coefficient, C_{loc} and graph signal filtering approaches where δ is the increased amplitude of the central point of the sphere.

Table 5.1: Percentages (%) for different methods in correctly locating spheroid centre (Centre) and in identifying spheroid at any point (Any) over all sizes of strength δ

Locate	max	C_{loc}	$\hat{\mathbf{W}}$	$\hat{\mathbf{W}}^3$	$\mathbf{L}$	$\mathbf{L}^3$	$e^{-\mathbf{L}}$	$e^{-3\mathbf{L}}$
Centre	9.7	<u>18.4</u>	17.4	1.9	4.8	2.5	1.8	16.6
Any	28.2	<u>41.1</u>	30.3	7.7	16.7	10.3	4.5	21.4

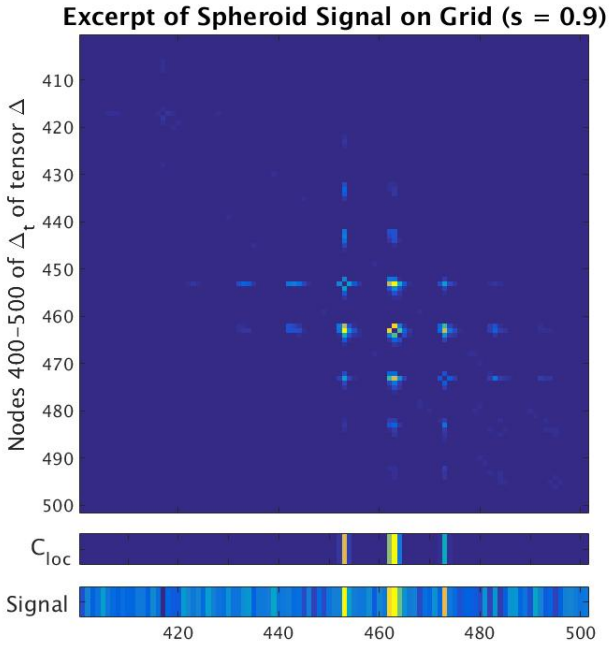

Figure 5.4: Example of reduced noise and increased accuracy through clustering coefficient. The highest amplitude is detected at node 452, however the maximum clustering coefficient, C_{loc}, detects the actual centre at node 462.

An example of how the proposed method is able to correctly identify a spheroid centre which has been incorrectly identified using the highest amplitude is shown in Fig 5.4. In this example, the increased amplitude of $3/4\delta$ given to one of the nearest vertices, 452, provides a larger overall amplitude to the δ given to the central node. By using the graph-variate method, however, this error due to noise is corrected since most of the nearest vertices to 452 have a very small comparative amplitude to those of the true centre at 462.

Analysis of the formulation of the C_{loc} shows its power for the suppression of noise and promotion of clustered phenomena. In the problem illustrated we can consider the expected value of the signal triple

$$\begin{aligned} &E[(x_i + x_j)(x_j + x_k)(x_k + x_i)] = 8E[x^3] \\ &= 8(\mu_3 + 3E[X]E[X^2] + 2(E[X])^3) \\ &= 8(\mu_3 + 3\mu(\sigma^2 + \mu^2) + 2\mu^3). \end{aligned} \tag{5.33}$$

where μ is the mean and μ_3 is the third moment of variable x. For only noisy data $x \sim$

$\mathcal{N}(0, \sigma^2)$, this is just zero from the fact that odd moments of a symmetric distribution are zero and $\mu = 0$. On the other hand, the expected value for $x \sim \mathcal{N}(\delta, \sigma^2)$, i.e. data with true value δ in the presence of noise, is $24\sigma^2\delta + 40\delta^3$. For the GSP filtering approaches, the adjacency matrix provides $E[x_i] = 0$ for noise and $E[x_i] = \delta$ for the true value which explains why it also fairs well at detecting the correct centre point, whereas the Laplacian provides $E[x_i - x_j]$ which is zero for noise and true value which explains its poor performance here. Of course, we note that this may be a very specific instance, there are undoubtedly other scenarios where the Laplacian/adjacency filter approach may fair much better. However, this highlights the necessity for the appropriate consideration of analysis for the problem at hand which can be assessed more fully within the proposed graph-variate framework. To increase comparability and the pursuit of a simple example, these approaches are chosen to be free from parameters and more complicated methodologies such as using iterative denoising. We recognise, though, that other more elaborate techniques such as implementing wavelets using a dictionary of spheroid shaped atoms [136] or joint time-graph denoising [83] may provide a more intensive treatment of the problem.

5.4.3 GVD connectivity of resting state EEG data

We study the eyes-closed, eyes-open dataset of 129-channel EEG activity as detailed in Section 4.3.1. These long recordings- 4.4355 $\pm$ 0.2861 mins (mean $\pm$ SD)- allow us to arbitrarily take windows starting at the 1000th sample (5s) to avoid the possibility of pre-processing artefacts at the beginning of the signal. We choose epochs, τ, lasting $16, 32, 64, \ldots, 2048$ samples (80ms up to 10.24s). We investigate dynamic connectivity using correlation, coherence and PLI in Alpha. For analysis, modules (subsets of nodes) of interest are chosen based on observable differences in the average weights over graphs computed from the largest window-2048, Fig 5.5. Choosing modules, instead of global connectivity, allows us to compute our phase-based methods without redundancy (5.29). Clearly, around 1-30 nodes and 60-90 nodes show differences in all connectivity measures (Fig 5.5, black lines mark 30, 60 and 90), thus we choose these as Module A and Module B, respectively, to compare our methods. Note that modules are chosen here from visual inspection as opposed to using module detection algorithms since such algorithms can be expected to choose different nodes between groups whereas we want a robust and unbiased comparison of modules between groups. As a result,

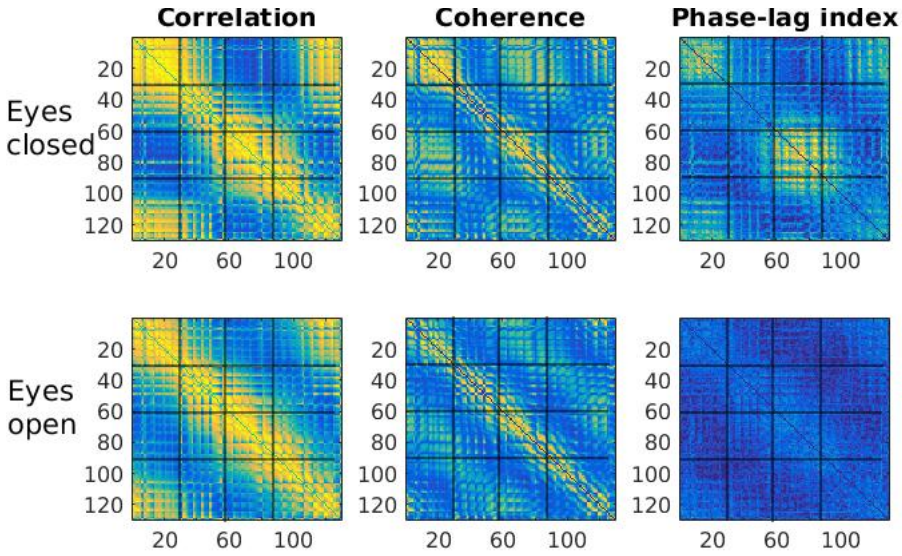

Figure 5.5: Weighted graph adjacency matrices of correlation, coherence and PLI for eyes closed (top) and eyes open (bottom) conditions. The colour axes, yellow being the largest weights, are the same for eyes open and eyes closed conditions. Modules are selected based on the most different activity between conditions- Module 1: nodes 1-30, Module 2: nodes 60-90, indicated by the black lines.

these modules are spatial, based on proximity of electrodes, rather than topological. Modular connectivity is computed as per the formula:

$$\sum_{t\in T}\sum_{i\in\mathcal{V}_a}\sum_{j\in\mathcal{V}} c_{ij}F(x_i(t),x_j(t)), \tag{5.34}$$

where $\mathcal{V}_x$ are the module vertices and T is the epoch of interest. Here, i sums over the module and j sums over the entire graph to assess the modules effects on the entire graph. Equation (5.34) is applied for correlation using (5.20) and (5.23), coherence using (5.25) and (5.26), and PLI using (5.28).

For this dataset we seek to clarify the usefulness of our methods compared to weighted graphs by themselves, as implemented in e.g. [73–76], as well as the benefit of the graph support in GVD connectivity as opposed to using un-weighted node space functions i.e. putting all weights equal 1 in (5.34), as used in [77].

For modules A and B, we compute GVD connectivity using the pair (T,τ), where the graph weight is computed over epoch τ and the node function over disjoint epochs of length T,

Table 5.2: Number of cases (T, τ) and p-values within those cases (cases:p-values) for which GVD connectivity (GVD) finds more significant differences ($>$) than node functions alone (NF) and vice versa. First column indicates GVD method used (graph/node function) where Cor-correlation, Ch- Coherence, sqd- squared difference, ico- instantaneous correlation and phs-sign of phase difference.

Method	**Module A**		**Module B**	
—	GVD>NF	NF>GVD	GVD>NF	NF>GVD
Cor/sqd	8:16	3:3	10:21	0:0
Cor/ico	4:4	3:3	6:11	0:0
Ch/sqd	14:62	0:0	6:7	1:1
Ch/ico	9:15	6:6	2:2	3:3
PLI/phs	10:21	4:5	4:9	6:9
Total	45:118	16:17	28:50	10:13

such that $T \leq \tau \in \{16, 32, 64, \ldots, 2048\}$. This gives a total of 36 cases corresponding to each combination of (T, τ) and a minimum of one p-value (when $T = \tau$) and maximum of 2048/16= 128 p-values for these cases. For each (T, τ) we then compute the density, $(\text{differences found})/(\text{total possible})$, of significant p-values found from paired t-tests of eyes closed vs eyes open conditions across the 16 participants. The results for each (T, τ) are shown in Fig 5.6 for modules A and B for GVD connectivity, the node functions by themselves (no graph) and a dynamic graph approach (graph only).

It is clear that for both modules the GVD connectivity approach performs better than the graph only approach for correlation and coherence. The PLI fairs poorly in this paradigm in general, but we shall see later that it may be leveraged to greater effect in time-locked task presentation data. It is not clear by observation whether the GVD connectivity approach is better than the no graph approach. To see this more evidently, we compute the number of cases, (T, τ), for which GVD connectivity outperforms the no graph approach and vice versa, as well as the greater number of significant p-values shown by GVD connectivity within those cases and vice versa. Table 5.2 shows the results. We see that GVD connectivity consistently outperforms the no graph approach with a total of 45 cases, consisting of 118 a p-values, in which it exceeds it in module A, and 28 cases, consisting of 50 p-values, in which it exceeds it in module B. The opposite, in which the no graph approach exceeds GVD connectivity is much lower with just 16 cases, consisting of 17 p-values, in module A and 10 cases, consisting of 13 p-values in module B.

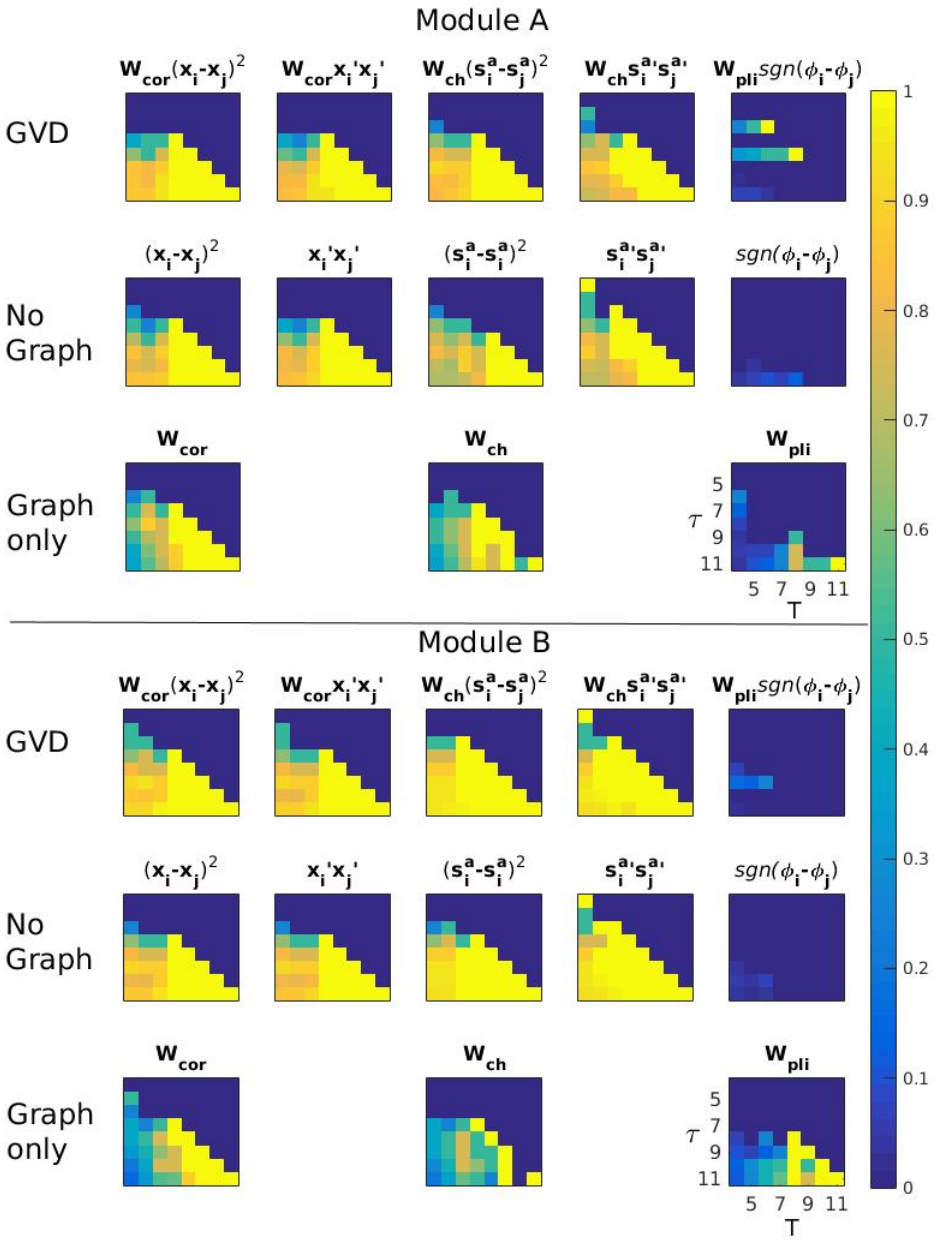

Figure 5.6: Results of eyes open vs eyes closed EEG data for Module A, top and Module B, bottom plotted by density of p values which are significant for $T < \tau$. GVD (first row) is GVD connectivity where the graph comes from τ and the GVD is computed over T. The axes of τ against T, shown on the bottom right plot, indicates the signal length considered in powers of 2, i.e. 5 is $2^5 = 32$, etc. No graph (second row) is the non-graph weighted node space function. Graph only (third row) refers to graphs computed over T. $\mathbf{W}_{cor}$ is the adjacency matrix of correlations, $\mathbf{W}_{ch}$ of coherence and $\mathbf{W}_{pli}$ of PLI. Here, $\mathbf{x}_i$ is the original signal, s_i^a the signal envelope and ϕ_i the instantaneous phase, where $\mathbf{x}_i'$ and $\mathbf{s}_i^{a\prime}$ are the signals minus their expected values as in (5.23) *and* (5.26)*, respectively.*

5.4.4 PLI during face recognition task

To try the PLI in a more appropriate task-related setting where consisting phase dependencies of brain function over many trials can be picked out, we look at a face presentation task, detailed in [137]. The dataset consists of 16 subjects undergoing a face presentation task lasting 1.5 seconds (0.5s pre-stimulus) down-sampled from 1kHz to 250Hz. Mean and SD of trials is 294.19 $\pm$ 2.32. After bandpassing in Alpha (8-13Hz), the PLI is computed for each trial and then averaged to construct an adjacency matrix per subject. GVD connectivity and the no graph approach are then conducted using sign of instantaneous phase differences. This is conducted per trial, then averaged over trials after which the absolute value is taken. Fig 5.7 shows the mean adjacency matrix over subjects ((a) top right) and the resulting C_{loc} for instantaneous phase and GVD connectivity estimates, averaged over subjects. In the GVD connectivity, we can see clearly a strong pattern of dynamic connectivity in nodes 40-60 occurring around 0.3-0.5s after stimulus which dies away and then appears to return again near the 1s mark. This activity occurs after N175 event related potential known to play an important role in face perception tasks [138], suggesting a post N175 phase-based functional response to the visual stimuli. Topoplots, where the sum of C_{loc} of node i is mapped as an intensity to the electrode position, confirm that this is more evident using the GVD approach, Fig 5.7 (b), where a strong polarity of activity from front right to back left from 0.3-0.5s reoccurring at 0.9-1s is contrasted with a drop in activity from left to back right. Activity from 0.3-0.5s is suggested also in the top left of instantaneous phase only but is less apparent and the reappearance near 1s is not evident.

5.4.5 ERP correlations of VSTM binding

Here we study the data described in Section 4.3.2. The 30 channels were re-referenced to the average EEG activity having already been processed using ICA to remove ocular artefacts. From the continuous EEG, we extracted epochs of 1.2 s starting at -200 ms pre-stimulus onset (baseline). We focused on the encoding (i.e. study display from 0-0.2 ms) and maintenance (delay from 0.2-1000 ms) periods of VSTM, since these seem to be the stages of memory informing about the functional principles of organisation with regard to capacity and format of representation (shape vs binding) [127]. We then computed the average Event Related Potential

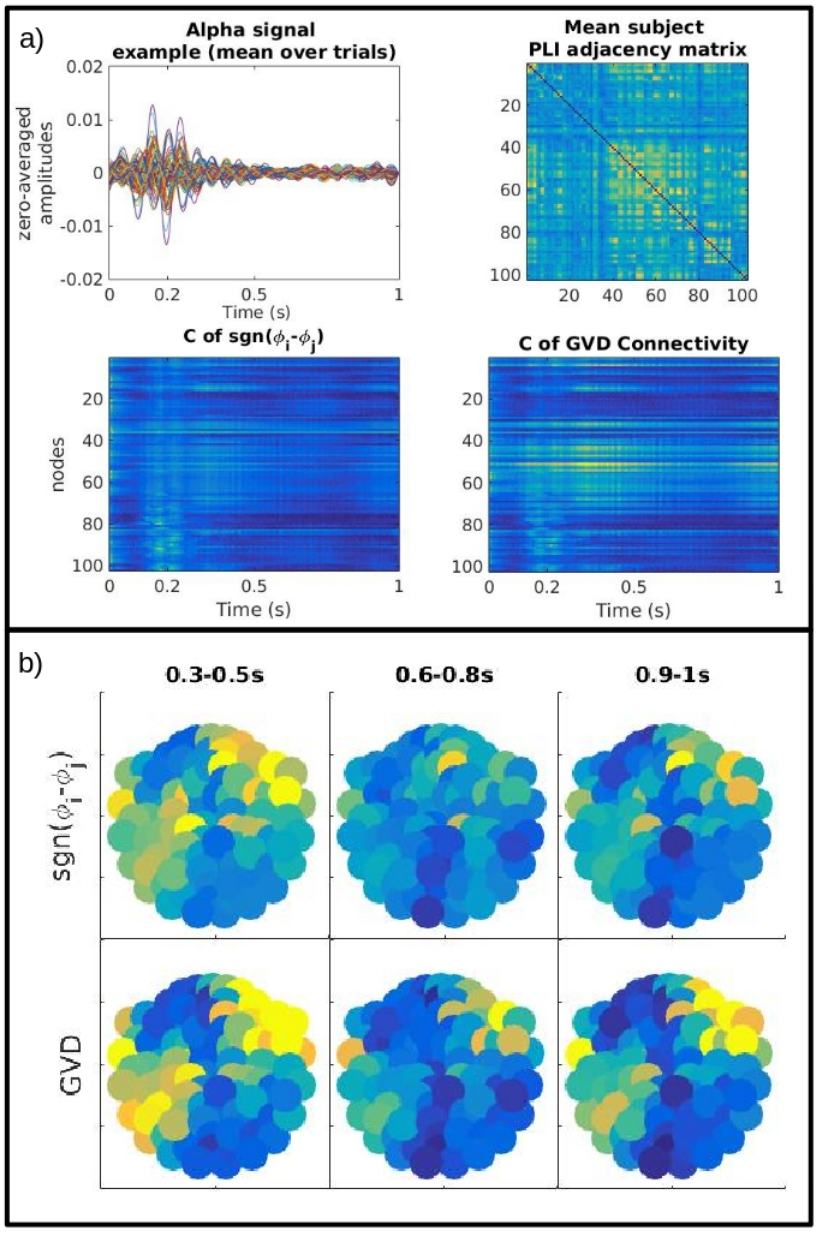

Figure 5.7: a) Phase activity from a face presentation task. Top left is the alpha signal for one subject. Top right is the mean connectivity adjacency matrix over all subjects. Bottom left is C_{loc} for each node at each time point for instantaneous PLI, averaged over subjects. Bottom right is C_{loc} for each node at each time point for GVD connectivity, averaged over subjects. b) Topoplots of the sum of C_{loc} phase activity in a given time window from a face presentation task. Colour axis has a minimum (blue) of the 10th and maximum (yellow) of the 90th percentile over all values, time points and subjects. Top is for instantaneous PLI and bottom for GVD connectivity, averaged over subjects.

(ERP) signal over correct trials (number of correct trials per participants, per condition: mean-65.7, SD- 9.27) for each VSTM condition performed by each participant resulting in a set of

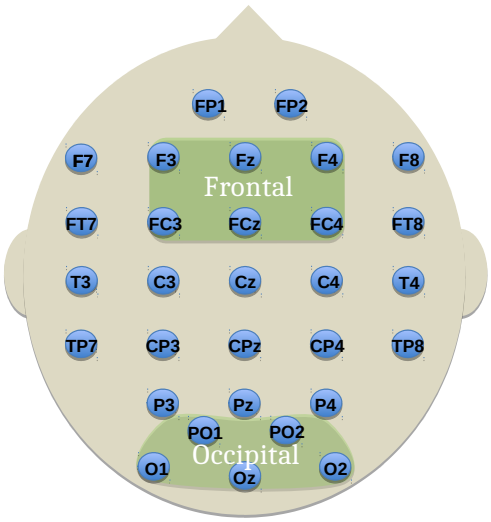

Figure 5.8: The frontal and occipital modules defined for network analysis. Labelled electrodes follow the 10-20 system.

4×19 thirty-channel EEG signals.

For each subject and task, graphs were created for encoding and maintenance periods where the edge weights were defined as the absolute value of correlation coefficient [120, 139, 140] between the pairwise channels for the broadband of frequencies (0.01-40Hz). The correlation coefficient is chosen, rather than phase-dependent connectivity, in order to analyse amplitude related effects from ERPs. The broadband was considered to reflect the real-time amplitudes important to our novel analysis and in keeping with our processing-light approach. Further, a previous study of ERP broadband analysis on these tasks gave promising results [128].

In order to find differences in cognitive task conditions that are representative of the sampled population (e.g. controlling for sources of individual variability such as head size, small electrode displacements, etc.), we considered activity over broader regions involving several electrodes, relating to modules of the constructed graphs. To this aim, we defined two modules using well known scalp regions (see Fig 5.8) which are relevant to working memory processing [141] and previously reported to be involved in the task investigated [127, 142]. These are the frontal module (F3, Fz, F4, FC3, FCz and FC4) and the occipital module (O1, Oz, O2, PO1 and PO2). To avoid combinatoric issues, these modules were chosen on physiological principles to be symmetric, of comparable size and with considerations of locality and generality in mind. Defining a priori modules on the graph allows us to combine these hypotheses with the

Table 5.3: Average node GVD connectivity over task conditions and participants for each electrode

FP1	-32.45	**F8**	-19.33	**T3**	-11.74	**CP3**	-10.27	**P4**	-30.00
FP2	-32.07	**FT7**	-15.5	**C3**	-10.36	**CPz**	-13.60	**O1**	-59.71
F7	-23.67	**FC3**	-17.66	**Cz**	-14.00	**CP4**	-16.01	**Oz**	-54.68
F3	-23.96	**FCz**	-20.21	**C4**	-12.81	**TP8**	-26.07	**O2**	-55.89
Fz	-28.09	**FC4**	-16.09	**T4**	-14.75	**P3**	-25.70	**PO1**	-56.29
F4	-20.32	**FT8**	-13.45	**TP7**	-17.45	**Pz**	-26.32	**PO2**	-46.65

information inside the signal, creating a topology of potential similarity between electrodes.

The composition of the modules (i.e. electrodes chosen) was constructed after considering the node GVD connectivity (5.15) computed for each node of the graph during the entire encoding and maintenance period to determine their suitability for the modules (i.e. using GVD connectivity to identify outliers). The occipital module was chosen considering all those electrodes in the occipital region. This choice is evidentially justified by the GVD node connectivity where there is clearly far larger magnitude values than in the rest of the electrode array (see Table 5.3, column 5). In choosing the frontal module we wish to consider a comparable size of module to the occipital. If we consider the suitability of electrodes FP1 and FP2 for the frontal module, we see these electrodes have node GVD connectivity over 2 SD above the mean drawn from the rest of the electrode array excluding the occipital electrodes which exhibit obviously stronger values (-18.56 $\pm$ 5.97 mean $\pm$ SD) thus we excluded them to avoid their overpowering influence since this highlights a strong contrast in activity. We then seek to form a symmetric module of comparable size to the occipital region, which leaves $\{$F3, Fz, F4, FC3, FCz and FC4$\}$ as the physiologically feasible choice (Fig 5.8). Normalisation to correct for such influences is neither obvious nor advisable since each edge in a graph corresponds to two nodes and any such process would act to negate the heterogeneous nature of the underlying EEG network degree distributions [1]. We investigated differences in the encoding period (0-200ms) and the maintenance period (200-1000ms) of the tasks by analysing the total modular weight:

$$w_{\mathcal{V}_a} = \sum_{i \in \mathcal{V}} \sum_{j \in \mathcal{V}_a} w_{ij}, \tag{5.35}$$

of the specified modules. We contrast these values for shape vs binding conditions in the left hemifield and in the right hemifield using paired t-tests. Using modular GVD connectivity

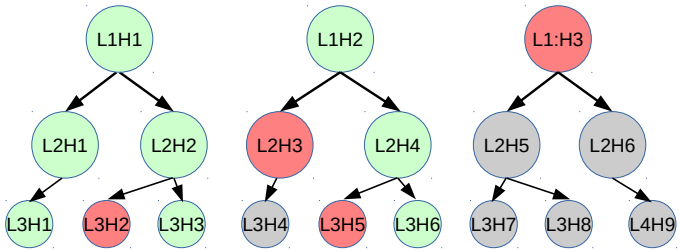

Figure 5.9: Example of hierarchical hypothesis tree for hierarchical false discovery rate procedure. 'LiHj' indicates the jth comparison in the ith level of the hierarchy. Red indicates no discovery, green indicates valid discovery, grey indicates exclusion from correction procedure due to false or no discovery made.

(5.16), we then introduce a second level of analysis to discover if particular parts of the original epochs are driving the discovered effects. Given the clear hierarchical structure of the hypotheses, we then use hierarchical False Discovery Rate (FDR) [143] to control for Type-I errors and allow a only the strongest temporal differences to be shown. Hierarchical FDR follows a level by level procedure of false discovery detection where a parent-child relationship is evident between these levels. Only those hypotheses whose parents were accepted as true discoveries are considered in the next level. In our study, the parent hypotheses relate to the total modular weights and the child hypotheses relate to GVD connectivity. Fig 5.9 shows a model of a hypothesis hierarchy and the principles of rejection and acceptance of discovery through the FDR corrective procedure. We implemented a strict FDR with $q = 0.05$ throughout the procedure [143].

From the effects found in the edge weight testing, we compute the modular GVD connectivity (5.16) for the frontal and occipital modules and the between module GVD connectivity (5.18) between the frontal and occipital modules. Note that, for module $\mathcal{V}_a$, the latter is contained in the former, i.e. $\theta_{\mathcal{V}_a,\mathcal{V}_b} \subset \theta_{\mathcal{V}_a}$, so that the latter probes the modules specifically for the interaction of the frontal and occipital modules.

We contrast shape and binding values throughout, therefore metrics are usefully presented as 'shape−binding' which implies the difference of the given metric values between the shape and binding condition. For this reason, we present box plots indicating where the 0 line is for the modular weights in the frontal and occipital modules (Fig 5.10, top left), and the GVD connectivity for modules and between modules (Fig 5.10, bottom row). A summary of

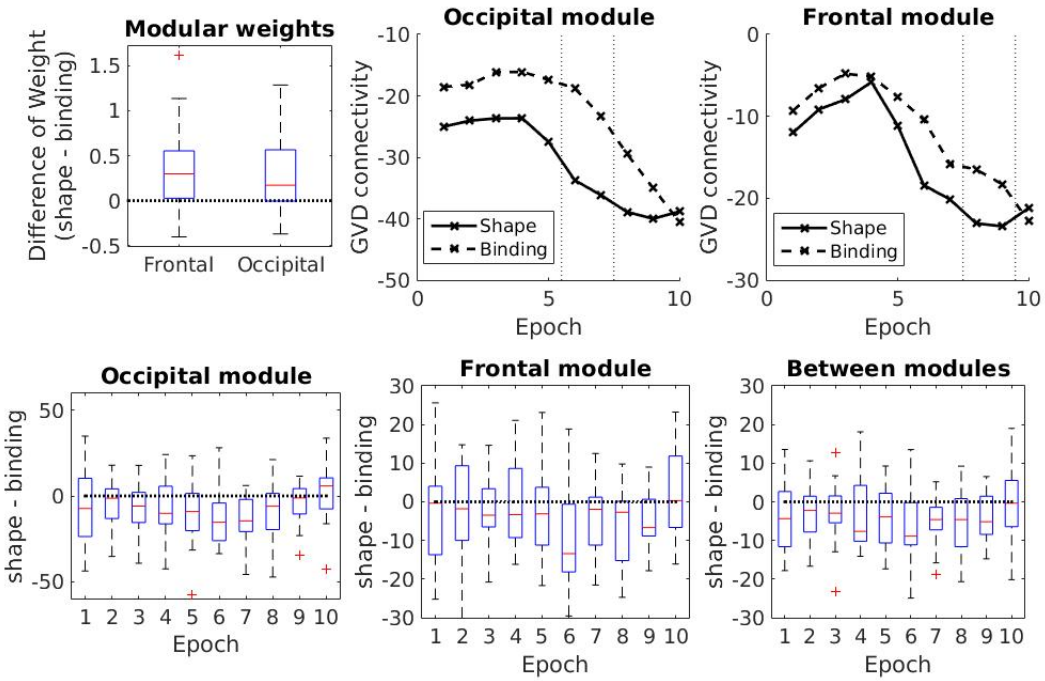

Figure 5.10: Top Left: difference in modular weights between the shape and binding tasks in the Frontal and Occipital modules. Top centre and right: Evolution (mean over subjects) of modular GVD connectivity of shape (solid) and binding (dashed) in the right hemifield during the encoding period for the occipital module (centre) and the frontal module (right) calculated over non-overlapping 20 ms (5 time sample) windows. The dotted lines indicate the beginning and ending of the epochs displaying significant differences in activity. Bottom: Box plots of modular GVD connectivity for each epoch for occipital (left), frontal (centre) and the between module GVD connectivity of Frontal and Occipital modules (right), where 'Epoch' refers to the 20ms windows, labelled consecutively '1' to '10'.

the results at two levels of analysis is presented in Tables 5.4 and 5.5. Paired t-tests were performed over participants for the measurements obtained for shape and binding conditions and subsequent p-values were controlled using hierarchical FDR [143]. The normality of the distributions was tested for each paired t-test using the one-sample Kolmogorov-Smirnov test. No significant deviations from the normal distribution were found at the 5% level. We report the following:

Level 1:

In the first level, the long-term total modular edge weights computed from the absolute values of correlation are analysed for the conditions to be contrasted. These contrast are left shape vs left binding and right shape vs right binding in both Frontal and Occipital modules during

both encoding and maintenance periods. Results found here thus inform on which periods, modules and task related hemifields are important in binding tasks. From the paired t-tests, after FDR correction, significant differences were found for the contrast involving shape vs binding conditions in the right hemifield (left hemisphere stimulation) for both Frontal and Occipital modules during the encoding period (see Table 5.4). These showed that the binding condition weights were less than those of the corresponding shape conditions (Fig 5.10), top left. No differences were found in the maintenance period and, further, no differences were discovered when contrasting shape vs binding conditions in the left hemifield for either encoding or maintenance periods. The supplementary material (section 3) of [4] provides a parallel analysis for modules chosen for different scalp areas, showing that we do not miss out on important activity happening elsewhere in the EEG signal correlates and further highlights the anatomical specificity of the activity elicited during this memory paradigm. We further consider an analysis of the sensitivity of electrode selection for our modules by removing and adding electrodes to the module to see how this effects the results. The results shows our methodological approach is robust to small modifications in electrode choice, demonstrating that the physiological considerations made for module choice do not substantially influence the results of the study.

Level 2:

In the second level of analysis, we perform modular GVD connectivity analysis over non-overlapping 20ms (5 time samples) windows over the modular weights (signed correlations) in Level 1. Our analysis now focuses only on those hypotheses from which their parent hypothesis in the first level were seen as 'true discoveries'. Thus, we present results of the GVD connectivity for Frontal and Occipital modules during the encoding period of shape vs binding condition contrasts displayed in the right hemifield. An extended table of results including those for the left hemifield can be found in the supplementary material (section 4) of [4]. Further, we study the GVD connectivity between Frontal and Occipital modules to discover if there are epochs where dependencies occurring between these regions show strong effects.

After FDR correction, effects are found in the GVD connectivity of the occipital module in the epochs between 100-120ms and 120-140ms, showing a larger negative GVD connectivity in

Table 5.4: p-values for paired t-tests of modular sum of edge weights in shape vs binding conditions. O = occipital module, F = frontal module, E = encoding period, M = maintenance period, L = left hemifield condition, R = right hemifield condition. Blue = true discovery, red = false discovery, black = null hypothesis not rejected at the 5% level.

O.E.L	O.M.L	O.E.R	O.M.R	F.E.L	F.M.L	F.E.R	F.M.R
0.1873	0.8709	0.0102	0.4514	0.2119	0.9040	0.0044	0.4806

Table 5.5: p-values for paired t-tests of Modular GVD (MGVD) connectivity and Between module GVD (BGVD) connectivity in shape vs binding conditions. Legend as in Table 5.4.

Time (ms)	MGVD - O.E.R	MGVD - F.E.R	BGVD - F.O.E.R
0-20	0.2036	0.4088	0.0942
20-40	0.0909	0.3891	0.0957
40-60	0.0432	0.1380	0.1408
60-80	0.0718	0.8074	0.1805
80-100	0.0254	0.1918	0.0412
100-120	0.0038	0.0465	0.0073
120-140	0.0010	0.0851	0.0028
140-160	0.0278	0.0070	0.0120
160-180	0.0919	0.0059	0.0167
180-200	0.6661	0.5464	0.9644

the shape task. In the Frontal module, an effect is found straight after this, between 140-160ms and 160-180ms, again showing a larger negative GVD connectivity in the shape task. Further, the GVD connectivity between modules shows an effect in the epochs of 100-120ms and 120-140ms (see Table 5.5). Notably, all the values in this study were strong negative values indicating generally matching information between the signals and the connectivity weights, as explained in the methods. This is exactly as is expected, since the signals are those from which the connectivity information is taken.

5.4.6 Discussion

The results from the simulations and resting-state and face presentation task EEG datasets show the greater power and suitability of implementing a graph-variate signal analysis approach over comparable state-of-the-art approaches. The greater power and flexibility, and thus suitability, to different datasets of the framework is displayed in Section 5.4.1, with instantaneous correlation outperforming Dirichlet energy, and Section 5.4.2, with GVD connectivity outperforming a wide range of GSP approaches. The generalisability also allows us to consider a wide array of problems in EEG functional connectivity– from resting-state

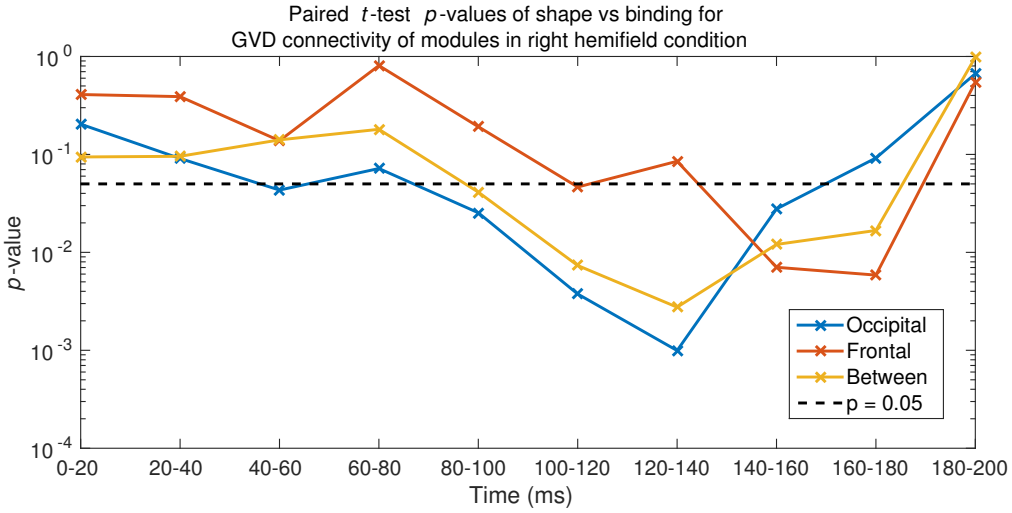

Figure 5.11: The p-values for shape vs binding contrasts in the right hemifield during the encoding period for GVD connectivity in the occipital module (blue), the frontal module (red), and between modules (yellow) calculated over non-overlapping 20ms (5 time sample) windows. The y-axis is on a logarithmic scale.

conditions in Section 5.4.3, to ERP analysis in Section 5.4.5, to phase-based task responses in Section 5.4.4.

In the ERP analysis of VSTM binding tasks, the results provide entirely new evidence of a focused prolonged functional difference between shape and binding conditions beginning in the occipital area at around 100ms, with a dependency between occipital and frontal areas from 100ms to 140ms and then shifting towards the frontal area between 140-180ms, see Fig 5.11. The strong chronological dependency of p-values over non-overlapping epochs is remarkable. Additionally, it is noticeable that all these effects have entirely vanished by the 180-200ms epoch, which is in accordance with the lack of findings found for the maintenance period.

From the GVD connectivity, we note that the effects reported in the occipital module appear to be driven by amplitude based activity between 100-140ms into the encoding phase of the task. This coincides with the P100 of visual evoked potentials and shows that with our methodology we are able to pick up on ERP activity over the network. During P100, the shape condition exhibited a noticeable dip in GVD connectivity which was much less apparent in the binding case (Fig 5.10, top centre). It is reasonable to suggest that this is caused by the greater work load in the binding condition. The involvements of visual association cortices in regions of the occipital lobe during short-term memory binding has been documented previously [127]. This

appears to be a key area of the visual integrative functions.

This evidence of contrasting brain function of shape vs binding conditions occurs during the encoding period in both the frontal and occipital modules supports the evidence found in Pietto *et al.* [128] that these tasks involve rapid functional activity in the frontal module which is picked up by the EEG. Recent electro-physiological studies indicate that frontal nodes may be contributing both specific (i.e. binding) and more general resources during working memory processing. The effect seen here between the frontal and occipital modules from 100-140ms and that seen in the frontal module between 140-180ms concurs with this, suggesting that a contrast exists in the functional dependency between these regions for shape vs binding conditions shortly after the onset of P100 activity. Further the activity occurring in the Frontal module indicates a difference in higher function post-visual processing.

As an important aside, these techniques are particularly timely for extending work done on uniting structure and function of the brain as in [144], in which the framework explores functional activations over the brain structure but does not yet accommodate for network relationships of synchronously active brain regions, for which graph-variate analysis provides a straightforward solution.

5.5 Conclusion

We defined and provided a general framework for graph-variate signals- unifying frameworks for multivariate signals and graphs. We developed novel analysis of graph-variate signals, considering general graph weighted node space functions. We showed the unique setting occupied by this new form of analysis within the framework, particularly with respect to graph signal processing. We then elaborated on novel methodologies for this analysis towards the temporo-topological analysis of multivariate signals and reliable connectivity estimation at signal resolution. In simulations we showed the robustness of the approach to finding correlations and detecting true activity within large datasets, in the latter instance outperforming similar state-of-the-art approaches. Pertinently, in differentiating coupling changes between EEG eyes-open and eyes-closed resting states, the methods generally outperformed graph only approaches. Using VSTM task data in healthy young volunteers, we also revealed dynamic task-related connectivity differences related to occipital and frontal

brain regions occurring within 80ms. Beyond brain networks, the implications of this theory reach into every domain in which network analysis of multivariate signals is used, such as in an application to gas-water two-phase flow [145].

Chapter 6

Illustration in Alzheimer's Disease

6.1 Introduction

Here we will apply the methods developed in Chapters 3, 4 and 5 to two clinical datasets of VSTM binding tasks of relevance to AD. Of note, these datasets concern different forms of AD, namely sporadic AD which, as its name suggests, occurs sporadically in the elderly population and familial AD which is hereditary, occurring with certainty during middle-age in those who possess the genetic variant [146]. It is particularly important to compare these two types of AD since findings from the rare familial form of AD are generally extrapolated to the common sporadic form of AD.

The VSTM paradigm we look at involves tasks which are specifically designed to test the ability of participants to store information of either shapes alone (shape) or shapes with associated colours (binding) for short-term memory recognition. In this instance it is found that patients are particularly impaired in the binding condition [70, 96, 97]. This VSTM task has been found to be both sensitive and specific to early AD [70] making it promising in the detection of preclinical disease [96, 97, 147]. Note, that these tasks are different from that detailed in Section 4.3.2 where, notably, the task is presented only in the centre of the screen, not in right and left hemifields, and the stimulus lasts 0.5s rather than 0.2s. Therefore, any results here cannot straightforwardly be reconciled with those already seen in Chapters 4 & 5 and we shall therefore treat these independently.

We shall assess hierarchical characteristics of EEG functional connectivity during VSTM binding of pre-clinical AD patients using degree variance (hierarchical spread) and

hierarchical complexity, of relevance to AD abnormalities as described in section 2.8, alongside an analysis of its GVD connectivity. Taking an EEG functional connectivity approach allows a non-invasive and practical analysis driven by the established hypothesis of functional disconnections behind AD pathology [148]. Due to the abnormal hub activity associated with discussed in Section 2.8, we expect to find abnormal behaviour in hierarchical characteristics of the network topology in patients compared to healthy control and due to the poor performance of patients in VSTM tasks we expect to find abnormal topological and temporal behaviour associated with the binding function. If evidence is found to support these expectations, it will establish interest for further research to be undertaken into the use of these approaches as early biomarkers of AD.

6.2 Methods

6.2.1 Visual short-term memory tasks

The binding function of VSTM is singled out by contrasting tasks for the recognition of coloured shapes, which requires binding of shape and colour in memory retention (binding), and the recognition of single shapes or colours which only requires the retention of constituent features. For reference, an illustration of the task is shown in Fig 6.1. In the change detection task assessment of VSTM for shape alone, the arrays consist of three different black shapes and in the binding task the arrays consist of three different shapes each with a different colour. Each task trial consists of an encoding period (500ms), during which a study array is displayed on screen, followed by an unfilled short delay (900ms) and test period with a test array. During the test period, participants are prompted to respond whether or not the objects in the two arrays are identical. The positions of the objects are randomised between arrays to avoid use of location as a memory cue. Both shapes and colours are chosen randomly for each trial from sets of eight shapes and eight colours. A randomly chosen fifty percent of the trials have the same objects in both arrays. In the other fifty percent, two shapes seen during the encoding periods are replaced with two new shapes selected from the set, whereas in the binding task two coloured shapes of the test display swap the colours they had during the encoding period. All participants start with a brief practice session before undergoing one hundred trials per task. Binding and shape tasks are delivered in a counterbalanced order across participants.

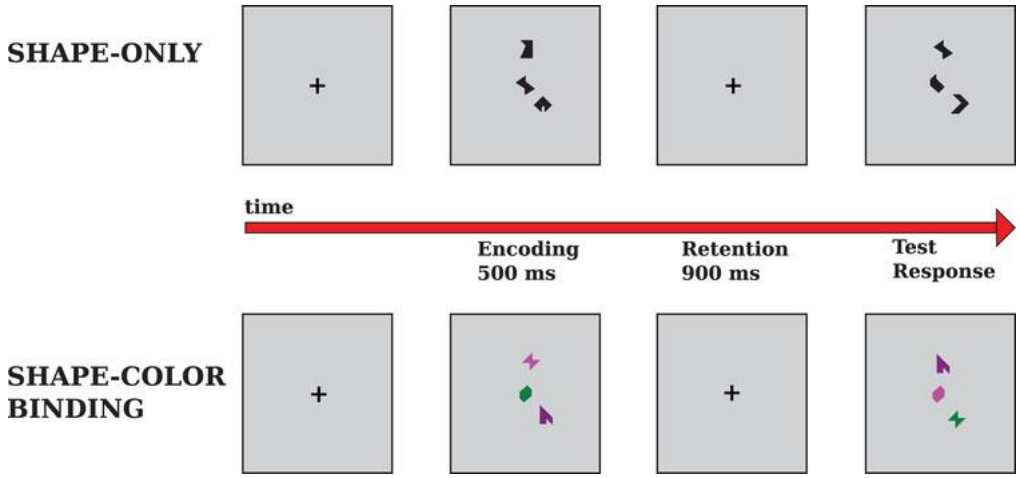

Figure 6.1: Illustration of chronology of the VSTM task presentation. Courtesy of Dr. Mario A. Parra [128].

6.2.2 Data

In this study, the aim is to test at the 'pre-clinical' phase of the disease to see if there is evidence of functional abnormalities of brain function that can aid in strategies for early detection. That is, the familial AD participants have not yet developed symptoms of AD, but will do so in due course. To test an equivalent sporadic AD group, we choose participants who are also in an asymptomatic, prodromal phase of the disease. Thus both populations are chosen as those with MCI.

6.2.2.1 Familial Alzheimer's disease dataset

The subjects are 10 patients– 44.4 ± 3.2 years old (mean $\pm$ SD), years of education 7.3 ± 4.1, MMSE scores 25.20 ± 4.50– and 10 healthy controls– 44.3 ± 5.6 years old, years of education 6.8 ± 2.9, MMSE scores 29.10 ± 1.10– from Antioquia, Colombia. Each patient carried mutation E280A of the presenilin-1 gene which guarantees early-onset familial AD. The data consist of sixty-channel EEG activity recorded with a sixty-four channel EEG cap using SynAmps 2.5 in Neuroscan at 500Hz and bandpass filtered from 1-100Hz with impedances below 10KΩ. Four ocular channels were discarded after being used to factor out oculomotor artefacts.

6.2.2.2 Sporadic Alzheimer's disease dataset

The subjects are 13 patients diagnosed with MCI– age 73.08 ± 9.01, education 14.08 ± 4.44, MMSE scores 26.46 ± 2.47– and 19 healthy controls– age 67.21 ± 10.14, education 16.50 ± 1.99, MMSE scores 29.50 ± 0.52– recruited from the Institute of Cognitive Neurology (INECO), Buenos Aires, Argentina. Criteria implemented for diagnosis derived from Petersen [149] and Winblad *et al.* [150]. Nine of the patients were at particularly high risk from AD conversion having been classified as single or multi-domain amnestic MCI while three classified as non-amnestic MCI multi-domain [151]. The data consist of one hundred and twenty eight-channel EEG activity recorded with a Biosemi Active 128-channel Two system at 512Hz and bandpass filtered from 1-100Hz. This was then down-sampled to 256Hz.

6.2.2.3 Notes for both datasets

Patients were evaluated with the MMSE and numerous other tasks with results detailed in [128]. The tasks were performed in an electrically shielded room with dim lighting. Subjects sat comfortably at a desk facing the task display screen. The subjects were checked to ensure that none had a history of psychiatric or neurological diseases. All participants provided written informed consent in agreement with the Helsinki declaration and the studies were approved by the Ethics Committees of the University of Antioquia and INECO.

We use data from the encoding period during the performance of shape only and shape-colour binding tests since deficits at this stage seem to be responsible for the VSTM binding problems found in AD [152]. This consists of 1.2s of continuous activity with 0.2s pre-stimulus. Signals were re-referenced to an average reference before proceeding. Further oculomotor artefacts were removed using visual inspection and independent component analysis and epochs with other artefacts exceeding $\pm 100 \mu$V were discarded. We seek to uncover underlying physiological substrates of the impaired binding function. In this way incorrect responses are not informative so only the trials where the subject responded correctly are included.

6.2.3 EEG functional connectivity estimation and binarisation

The PLI was computed to assess the phase-dependent functional connectivity of the EEG channels. The PLI is computed for each trial and for each signal pair after being bandpassed in Delta (0-4Hz), Theta (4-8Hz), Alpha (8-13Hz), Beta (13-32Hz) and Gamma (32-60Hz) using an order 70 FIR filter with Hamming windows. These connectivity computations are then averaged over trials for each task and for each subject to remove inter-trial variability and so better bring out the specific task function. The resulting averages constitute adjacency matrices of weighted networks, one for each subject-task-frequency band triple.

6.2.4 Characteristics of a network hierarchy

Before studying the network hierarchies, the weighted PLI connectivity networks are binarised using the CST which is known to provide a sensitive and powerful binarisation of EEG PLI connectivity [2]. The hierarchy of a network is here defined based on the number of edges adjacent to each node. Nodes with more adjacent edges are higher in the node hierarchy, being more central to the network topology. We study two indices of network hierarchies described in Chapter 3. The degree variance, V, measures the spread of the hierarchy and thus is an important indicator of the dominance of hub nodes. The hierarchical complexity, R, measures the complexity of interactions between hierarchy levels which provides a more nuanced evaluation of hierarchical order.

6.2.5 Graph-variate Dynamic Connectivity

We study graph-variate dynamic connectivity of PLI with node function (5.28), following the protocol laid out in Section 5.4.4. First, the average PLI matrix over trials is computed for each task of each subject. Then GVD connectivity is computed for each trial using the average PLI matrix and these results are then averaged to provide the mean GVD connectivity for each task of each subject.

6.2.6 Statistical tests

Differences of network index values for binding and shape are computed. These differences are contrasted between patients and controls and between elderly and middle-aged adults using population t-tests with statistical significance noted at the standard $\alpha = 0.05$ level. Effect sizes using Cohen's d for paired data are reported for significant differences.

6.3 Results

6.3.1 Observations of data

Initial observations of the data are shown in Fig 6.2 and 6.3 which displays the median PLI adjacency matrix over subjects in all 5 standard frequency bands for familial AD and sporadic AD datasets, respectively. It is notable that activity in Alpha in familial AD is remarkably similar to that of Theta in sporadic AD with a strong contrasting activity appearing smoothly over the matrix and that activity in Beta in familial AD also behaves similarly to that of Alpha in sporadic AD with patients showing increased activity in horizontal/vertical lines indicating hub activity (the specifics of nodes cannot be gathered here because of differences in the EEG layout). This suggests that relevant phase dependencies of these tasks corresponds to a drop in frequency which can be interpreted as age-related slowing of activity. There is also corruption apparent in Gamma of the familial AD dataset, so we will not use this band.

Fig 6.4 shows the adjacency matrix to the left of a connectivity map of the mean strength of PLI connectivity for each electrode in the scalp-space. For the connectivity maps, intensity is relative to the strongest (pure red) and weakest (pure blue) node values over patient's shape and binding performance and control's shape and binding performance to ensure comparability. These are shown for Alpha and Beta of the familial AD dataset (top and bottom left, respectively) and Theta and Alpha of the sporadic AD dataset (top and bottom right, respectively).

Alpha activity during tasks in the familial AD dataset (top left) shows strong frontal and occipital connectivity with weaker connectivity in the central region. Average patient connectivity appears weaker than average control connectivity and this is particularly apparent

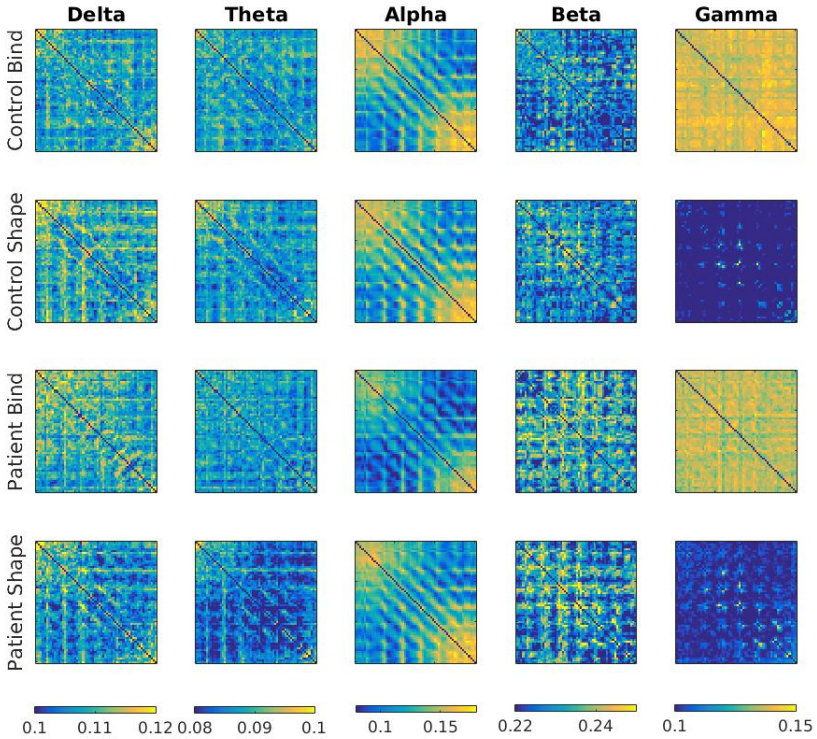

Figure 6.2: The median weighted adjacency matrices over subjects of PLI in 60 node familial Alzheimer's disease dataset.

in binding, although these are not statistically significant across subjects. Beta (bottom left) shows a multi-faceted contrast with connectivity being strongest in right and left temporal regions where average patient connectivity is notably stronger than average control connectivity, particularly in binding.

Moving to the sporadic AD dataset, we notice that the Alpha response in the younger familial AD and controls is mirrored in the lower frequency Theta band in the elderly. This supports the hypothesis of a slowed functionally related electromagnetic pulses with age in visual short-term memory, where functional phase differences at 8-13Hz in 50-60 year olds is occurring at roughly half the frequency, 4-8Hz, in the elderly. We thus use Theta in the sporadic dataset as a parallel to Alpha in the familial dataset. Similarly, in Alpha of the sporadic dataset we are obtaining connectivity maps more similar to Beta in the familial dataset with connectivity peaking in right and left temporal regions which seems particularly strong in patient binding. Thus we hypothesise that phase differences in Beta (13-32Hz) in 50-60 year olds is similarly

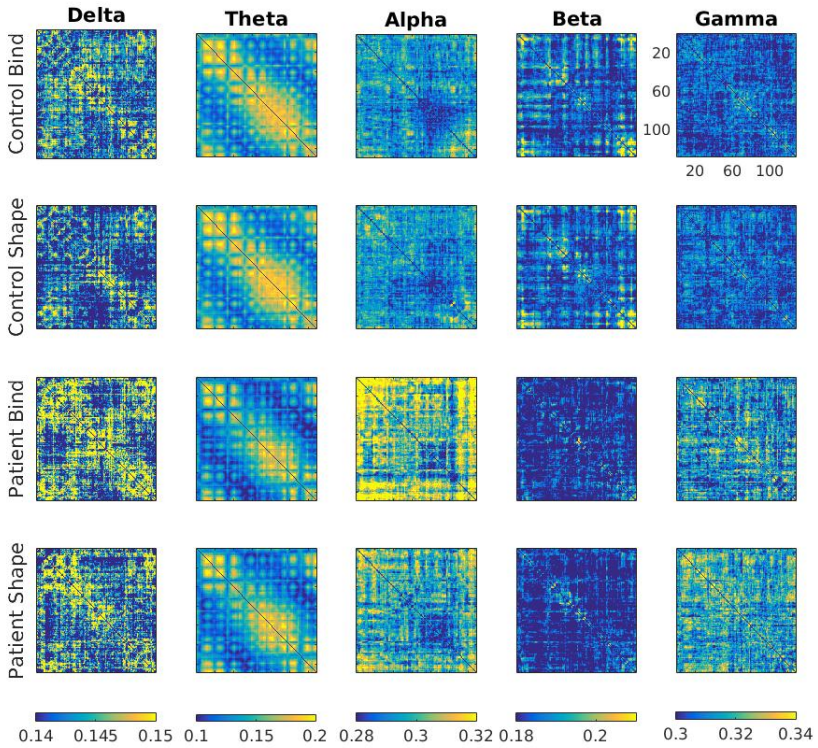

Figure 6.3: The median weighted adjacency matrices over subjects of PLI in 128 node sporadic Alzheimer's disease dataset.

slowed and should thus be evident in Alpha (8-13Hz, roughly half the frequency) in the elderly.

6.3.2 Hierarchical characteristics

For the familial AD dataset we compute V and R for all subjects in both tasks in Alpha and Beta. Population t-tests are then implemented for patients versus controls in the binding and shape tasks separately and paired t-tests are implemented for shape versus binding tasks in patients and healthy controls separately. This allows us to directly probe differences in MCI from healthy control in the binding task compared to the shape task and abnormalities of VSTM binding function in MCI from task differences compared to healthy control. A significant difference ($p = 0.0018$) is noted in V in Beta for shape versus binding in MCI, Table 6.1, left, with no difference found in healthy controls. In nine of the ten patients, the hierarchical spread of the Binding condition is larger than in the Shape condition, Fig 6.5, left, with an effect size of 1.3778. On the other hand, 7 out of ten controls exhibited an opposite characterisation.

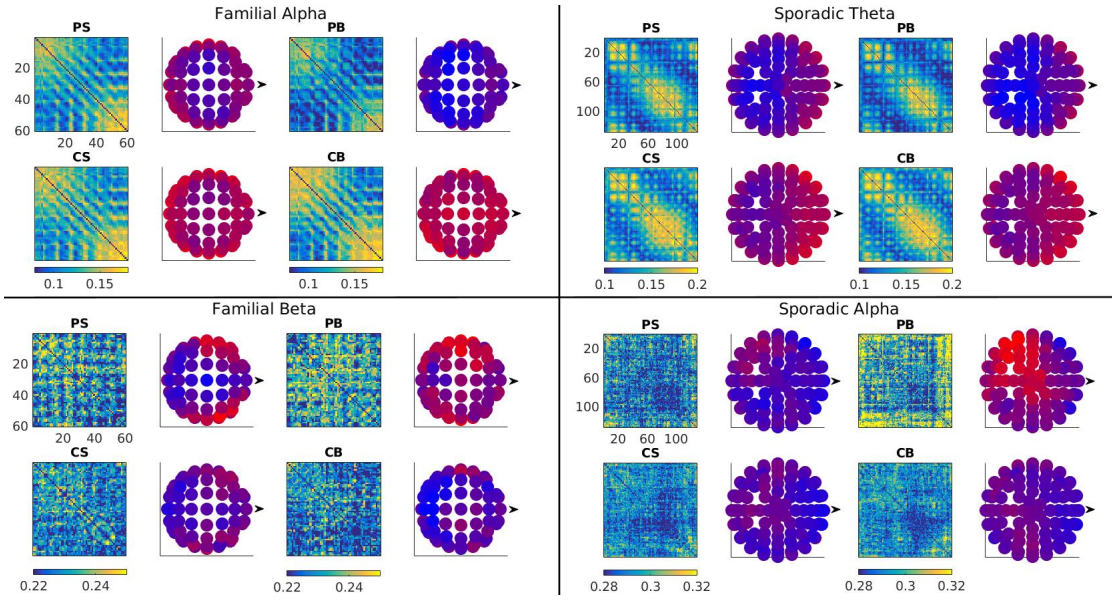

Figure 6.4: The median scalp-space EEG node functional connectivity over subjects of phase-lag indices. Top shows Theta and Alpha in the sporadic Alzheimer's disease dataset. Arrows indicate the forward facing direction. PS- patient shape, PB- patient binding, CS- control shape, CB- control binding.

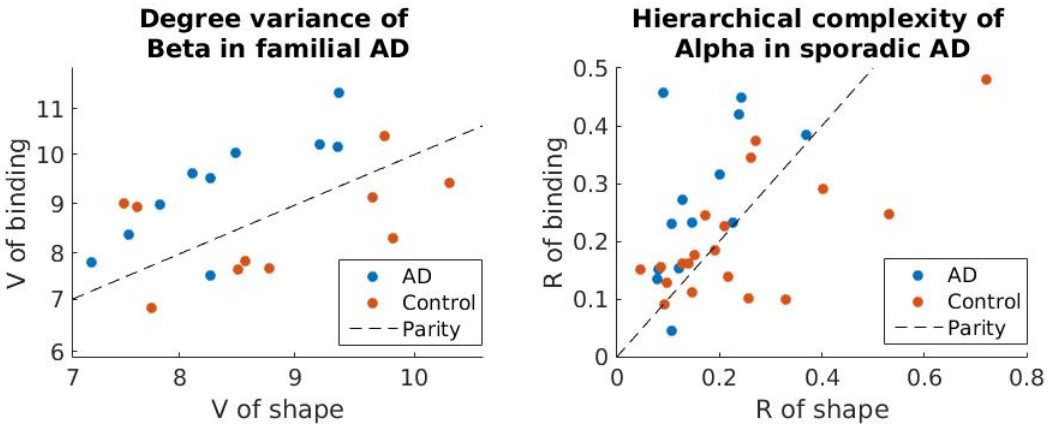

Figure 6.5: Scatter plots showing changes of hierarchical topology in phase dependencies of EEG activity in familial (left) and sporadic (right) AD patients in CST PLI networks.

Correspondingly, for the sporadic AD dataset we compute V and R for all subjects in both tasks in Theta and Alpha and population t-tests are then implemented for patients versus controls and paired t-tests for shape versus binding. A significant difference ($p = 0.0051$) is noted in R in Alpha, Table 6.1, right. In twelve of the thirteen patients, the hierarchical complexity of the binding condition is larger than in the shape condition, Fig 6.5, right, with an effect size of 0.9476. The controls are roughly balanced between higher shape and higher binding. A less

Table 6.1: Significant differences of hierarchical characteristics of CST PLI networks in familial AD. Presented are the p-values for population t-tests of shape vs binding for MCI and healthy controls and population t-tests of MCI vs healthy controls for shape and binding. Legend: MCI = MCI binding vs shape; CON = control binding vs shape; SHP = shape MCI vs control; BND = binding MCI vs control.

	Familial Alpha		**Familial Beta**		**Sporadic Theta**		**Sporadic Alpha**	
Test	V	R	V	R	V	R	V	R
MCI	0.9407	0.9785	<u>0.0018</u>	0.1385	0.8645	0.6725	0.0506	<u>0.0051</u>
CON	0.0697	0.8430	0.4059	0.2716	0.1927	0.3192	0.2686	0.2810
BND	0.0621	0.4453	0.1128	0.4303	0.2651	0.5123	0.5452	0.1405
SHP	0.4852	0.5389	0.2598	0.2880	<u>0.0293</u>	0.0714	0.8219	0.1745

convincing significant difference ($p = 0.0293$) is also noted in V between MCI and healthy controls during the shape task in Alpha.

6.3.3 GVD connectivity

We wish to check for temporal abnormalities of MCI in binding tasks during Beta in the middle-aged familial AD and Alpha in elderly sporadic AD, i.e. those exhibiting diferences in Section 6.3.2. The median over subjects for mean clustering coefficient of GVD connectivity at each node is displayed in Fig 6.6. Similarities between datasets are seen in strong activity occurring laterally over nodes for two small epochs with a dip in activity in between indicating that this phase-based activity plays an important role in functional processing of VSTM tasks. However, surprisingly, these epochs are not positioned at the same times over the datasets. The first epoch of increased activity in the familial AD dataset occurs between 0.16-0.24s and the second between 0.64-0.72s, whereas for the sporadic AD dataset they occur earlier, between 0-0.2s and 0.39-0.47s, respectively. Though it is reasonable to suggest that this sequence of activities relates to the same functional processing, the difference is surprising not least because the faster response appears to occur in the older people at lower frequencies. It is thus quite difficult to reconcile this information with the previous results in an obvious way, but we will carry on with statistical analysis of these epochs regardless.

Similarly as for the analysis of hierarchical characteristics, we seek differences between MCI and healthy controls in binding tasks and differences between binding and shape tasks in MCI to explore abnormal functioning behind the depreciated performance of binding in AD. Table 6.2 provides the results from population t-tests and paired t-tests, where appropriate, of activity

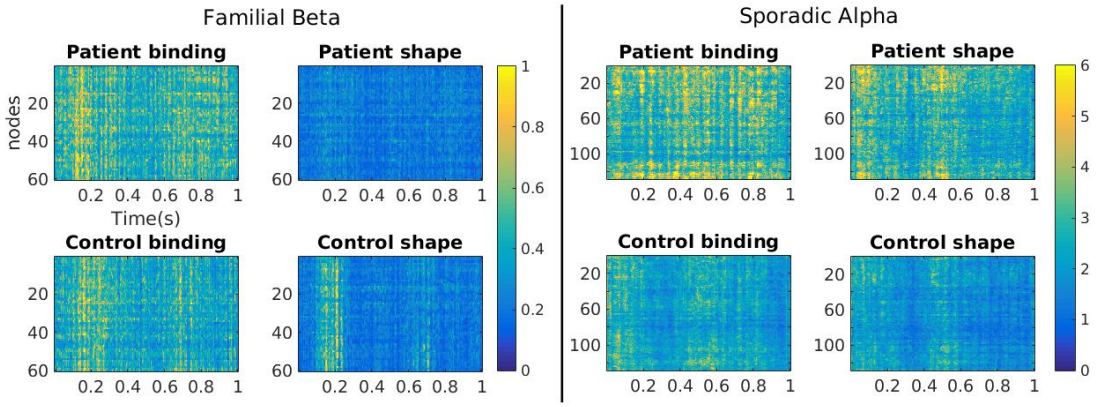

Figure 6.6: Median clustering coefficient over subjects of GVD connectivity. This is computed using PLI from EEG recordings of VSTM tasks performed by familial and sporadic MCI and healthy age matched control. Axes are as labelled on the top left most plot and frequency band and dataset are as noted in the panel titles.

Table 6.2: The p-values for paired t-tests of shape vs binding for MCI and healthy controls and population t-tests of MCI vs healthy controls for shape and binding. Values come from the mean clustered GVD connectivity for EEG PLI connectivity in the given epochs.

	Epoch (s)	**MCI**	**Control**	**Shape**	**Bind**
Familial AD Beta	**0.16-0.24**	0.0166	0.6029	0.0671	0.7526
	0.40-0.60	0.0169	0.0204	0.9290	0.9147
	0.64-0.72	0.0036	0.2564	0.1921	0.4584
Sporadic AD Alpha	**0.00-0.20**	0.8191	0.5901	0.6715	0.2127
	0.20-0.39	0.3952	0.3980	0.7906	0.0074
	0.39-0.47	0.6579	0.9405	0.2892	0.0998

in the aforementioned epochs as well as in the epoch of decreased activity in between. In the familial AD dataset, differences are apparent in MCI between shape and bind tasks in all three epochs, most convincingly ($p = 0.0036$) between 0.64-0.72s where binding is found to be greater than shape with an effect size of 1.2364. A difference in controls is only observed ($p = 0.0204$) in the middle epoch, 0.4-0.6s, again suggesting binding greater than shape with an effect size of 0.8884. No differences are found between MCI and healthy controls

In the sporadic AD dataset, a single difference ($p = 0.0074$) is apparent in the middle epoch between MCI and healthy control in the binding task. This shows greater activity in patients with an effect size of 0.9317. No differences are found in the shape task and no differences found between shape and binding in either MCI or healthy controls.

6.4 Discussion

The results of hierarchical topology analysis suggest that familial MCI patients with a certainty of progression to AD show a greater recruitment of hub activity in VSTM binding tasks than in shape tasks which may be explained by compensatory effects due to degradation of functional connections caused by AD. In elderly MCI patients at high risk of developing sporadic AD, the difference is more subtle and seen in a greater complexity of hierarchical organisation of functional connectivity, again suggesting possible compensatory effects due to normal network degradation. That these are both tested at MCI stage and it is not expected that familial AD lasts longer than sporadic AD means that the more obvious hierarchical degradation of familial MCI is unlikely to be due to more damage, for example. Together, the results suggest hierarchical characteristics are of high relevance in exploring functional EEG networks in early stage AD and show promise for developing potential biomarkers of disease.

Analysis using GVD connectivity provides interesting additional information about specific epochs of time showing abnormal activity in VSTM binding of MCI patients. In familial MCI, patients showed particularly increased activity during VSTM binding at 0.64-0.72s post-stimulus onset compared to the VSTM shape only tasks which was not present in controls. In this epoch the task generally elicits increased activity and so the greater activity in MCI supports the hypothesis of compensatory effects in binding function due to a degradation of the neural connectivity. These were not seen in sporadic MCI, but instead a mutually exclusive effect was seen between patients and healthy control during the binding function at 0.2-0.39s. This epoch of the task generally elicits a decrease in activity between epochs of heightened activity. The inability of sporadic MCI to settle to normal levels during this epoch suggests an increased functional difficulty in dealing with the binding task.

Given the data, the results from both of these forms of analysis could be interpreted in two ways. Either age related differences in the aberrant topology of functional EEG networks in early stage AD or early stage differences in how sporadic and familial AD effect the functional activity. Either of these hypotheses are interesting to explore and require a greater reference to specialist literature than is within the scope of this thesis. That being said, the promising results suggest that hierarchical complexity and GVD connectivity of EEG during VSTM tasks could

be progressed as a potential diagnostic tool for AD. Future studies will attempt to establish this and the hope is that such analysis could eventually be implemented as part of general screening assessments in the clinic. It is important to recognise that confounding factors in comparing network topologies of these datasets may exist relating to e.g. MMSE scores [37] or discrepancies in years of education and that more work needs to be done in order to understand these factors in future research.

6.5 Conclusion

Hierarchical characteristics of PLI EEG networks and PLI GVD connectivity were studied in familial and sporadic MCI subjects performing a VSTM binding task. Abnormal activity during binding was found in both datasets using both analyses. Particularly, it was found that familial MCI requires increased hierarchical spread in VSTM binding than in the VSTM shape task (Beta) whereas sporadic MCI requires increased hierarchical complexity in VSTM binding than in the VSTM shape task (Alpha), where the corresponding frequency bands containing these differences were found to relate to the similar task processing explained by slowed activity due to ageing. Further, GVD connectivity of VSTM binding was larger than VSTM shape in familial MCI whereas VSTM binding of sporadic MCI was larger than healthy controls. This showcases the fruitfulness of pursuing the novel analytical methods introduced in this thesis to a major societal problem.

Chapter 7

Conclusions and Future Work

7.1 Discussion

In this thesis, I introduced a range of novel complex network methods geared towards neuroimaging applications to understand more about brain function of which I focused on non-invasive EEG for its high temporal resolution and practicality. These were successfully applied to the problem of detecting early stage AD.

In Chapter 3, hierarchical complexity of network topology was defined on network degrees and a novel index to measure it was introduced based on the degrees of nodes' neighbours. Complementing this, the WCH network model was introduced, opening a new complexity paradigm between random and strict degree hierarchical topology. I showed that EEG functional connectivity, as measured by PLI, was hierarchically complex compared to a range of network archetypes and matched the maximum complexity achieved by the WCH model. The complexity of EEG topology was much better explained through this new paradigm than the pre-existing small-world paradigm, eliciting values that better approached those of EEG networks at all densities. I believe this should pave the way for a paradigm shift in how complexity of network topology is to be viewed.

Alongside this, the work covered the analysis of index curves [37, 38] (indices applied to each integer percentage binarisation of a weighted network) by suggesting simplifications of the integration/segregation concept and proposing a known index, degree variance, to characterise scale-freeness or, perhaps more accurately, the hierarchical spread of the degree distribution. I also introduced methods to generalise known binary network archetypes to weighted form

by evolving the archetype up to greater densities and eventual completeness before defining a complete weighted graph by taking an average over these densities. These developments allowed the comparison of the introduced weighted model against known archetypes over the entire possible density range.

In Chapter 4, our attention was then turned towards the important open problem of network binarisation [30,44,47,50,111,118–121] and how hierarchical complexity of network topology expresses itself in a binarised network. Included in this was a novel binarisation method- the CST- which binarises the network in medium density ranges (40-50%). Although current trends promote sparse densities, it was hypothesised that complex hierarchical topologies would be best captured by non-sparse densities. This hypothesis was tested by attempting to discern differences between populations of subtly different topologies defined using the WCH model. The evidence indeed backed the hypothesis, showing the power of using medium density ranges, including the CST, in network binarisation compared to weighted networks alone or using sparse methods such as MSTs or the ECO threshold. It was further shown that the CST was robust to random and targeted network attacks. This evidence was validated in practice using analysis of PLI connectivity from three EEG datasets in which the CST found interpretable differences in each case, where the MST, USP, ECO and weighted networks were found lacking.

In Chapter 5, the novel framework of graph-variate signal analysis was presented. This framework was inspired by research conducted in GSP in which a signal is supported over a graph topology. It was found that the GSP framework was too rigid for the direction I wished to explore and thus a generalised framework was presented to unify frameworks for multivariate signals and graphs. From this new setting, a new form of dynamic connectivity- GVD connectivity- was formulated with its main benefit being robust connectivity estimates at arbitrarily high temporal precision. Appropriate functions for GVD connectivity were provided for correlation, coherence and PLI, and a network science of GVD connectivity was illustrated. The proposed methods could detect evidence of a single correlated source within $128 \times 127/2 = 8128$ edges. They were then compared favourably to existing GSP approaches for a randomly travelling spheroid detection problem. Applied to EEG functional connectivity, GVD connectivity outperformed state-of-the-art dynamic connectivity approaches in

determining differences between eyes-closed and eyes-open resting states and were shown to be of promising use to detect important changes in task-related activity.

Finally in Chapter 6, I applied indices of degree hierarchy (degree variance and hierarchical complexity) and GVD connectivity to get new insights into aberrant topological and temporal connectivity behaviour produced in early-stage AD of both familial and sporadic types. These datasets involved performance of cognitive VSTM memory tasks (shape vs binding) which show promising potential as biomarkers of disease and it is hoped efforts to understand physiological underpinnings of these tasks in health and disease can help in detecting early signs when the disease has yet to show symptoms. The analysis revealed abnormal hierarchical topology in both sporadic and familial MCI. It also revealed that the form of abnormality for these two types was different- that binding function in middle-aged familial MCI required greater hierarchical spread than shape only whereas binding function in elderly sporadic MCI required greater hierarchical complexity than shape only. Healthy controls showed no differences. The GVD connectivity analysis revealed heightened PLI activity in binding of MCI patients in short epochs of task related function. Familial MCI showed a particular difference between VSTM shape and binding between 0.64-0.72s whereas the VSTM binding showed greater activity in sporadic MCI than healthy controls between 0.2-0.39s. This novel information opens up important considerations for clinical interpretation in terms of how familial and sporadic AD appear to disrupt normal hierarchical organisation of VSTM binding functionality in different ways, though both pointing to compensatory effects– increased hub activity vs increased complexity. Further, the GVD connectivity of PLI reveals temporally specific compensatory effects in binding as well as an unintuitive task based response where the older patients exhibit the same activity in lower frequencies and yet task-based response appears to occur more quickly.

7.2 Limitations

Certain limitations occurring in this thesis relate to the sample sizes of the real EEG datasets used, where we generally consider the comparison of populations of size 20 or less. Using larger populations would allow more powerful statistical testing and more convincing results. This is by no means unique to this thesis as obtaining data from large populations of

pathological patients is practically very difficult. In light of this, where possible extensive simulations have been implemented to provide more conclusive evidence for the benefits of these methodologies.

Regarding neuroscience imaging studies, it should be noted that, without a ground truth, we rely on the assumption that contrasting conditions provokes a contrast in functional connectivity. Also, although, the PLI was found to be reliable and straightforward to interpret it should be recognised that finding appropriate connectivity measures is a much debated topic with many considerations including hypotheses of how brain function takes place; the part and frequencies of the signals that should be used for a given paradigm; whether the measure should provide directed-ness; and whether the signals should be orthogonalised or relocated to the source space.

Regarding the VSTM familial and sporadic AD data, the study lacks a longitudinal component. Particularly, following up on the sporadic MCI patients to find out exactly those who go on to develop AD and also to assess cognitive effects as disease progresses in both datasets would significantly enrich the study. That being said, such longitudinal data are difficult to obtain due to problems of drop-out rates due to unexpected circumstances. In the future, it is hoped such data will become available and the first steps towards understanding VSTM deficits in disease progression can be undertaken.

7.3 Future Work

There are many routes for future work based on the contributions of this thesis owing to the introduction of various brand new methods. Notably, the abstract mathematical nature of the proposed hierarchical complexity and graph-variate signal analysis allows these methods to be applied in various neuroscience studies, but also to various other engineering problems of complex systems such as in economics [153], social sciences, fluid dynamics [145] and geophysics [3], with suitable adaptations. But I also hope these methods will inspire new ideas and further explorations into the theory of network science and multivariate signals and will expend most effort on these considerations for the rest of the chapter.

Hierarchical complexity provides an interesting nuanced analysis of network topology and

shows success in distinguishing pathological from healthy brain function. However it appears rather unpredictable compared to other more straightforward indices, indicated by the higher within population variability of the index compared to others, Fig. 3.3. I conjecture that this is likely owed to the fact that the analysis within neighbourhood degree sequences of the same length may be easily influenced by small changes in topology. Taking a careful analytical consideration of this would hopefully lead to an index more predictable within populations and studying neighbourhood degree sequences in other ways, for example the number of unique degree sequences in the network, is also of interest.

The WCH model is able to mimic well the topological characteristics of EEG networks and has parameters which can provide powerful simulation abilities usable for the assessment of new methods such as in binarisation techniques (as shown in chapter 4) or network indices. There are several apparent ways I consider to improve or extend the model. The first would be to seek a model with fewer parameters. Particularly, the idea of hierarchy levels has some benefits, but we would expect the reality of a real world hierarchy to not be so rigid and pursuing, for instance, statistical distributions to determine the hierarchy is of high interest. Secondly, the initial weights on which the model is founded are uniformly random, whereas real world networks are likely to have more nuanced preliminary foundations for connections, such as, for example, the physical space in which the network is set. This is obviously the case for the brain where the the neurons are physically inter-connected via the axons and dendrites. Thirdly, one must consider the form with which a known hierarchy acts on the foundational weights. The model adds hierarchical weights linearly to the foundational weights, but it is worth considering whether other transformations (e.g. multiplication) are more effective. These aspects will be explored in a future work currently under preparation.

The CST provides a medium density binarisation of the network which was shown to be more useful for EEG PLI networks than weighted networks and MST, USP and ECO thresholds. However, it is of interest to consider if complex hierarchy is of paramount importance to the utility of a binarisation method, in which case thresholds based specifically on such a concept would be preferred.

The framework of graph-variate signal analysis provokes fundamental considerations of how to formulate the best analyses of multivariate signals using graphs for any given problem. The

methods of GVD connectivity and GVD network science provide powerful and interesting tools to analyse data in new ways and I have already established a collaboration with a group at Tianjin University, Tianjin, China, on applying basic GVD connectivity methods to a gas-water two-phase flow problem which now has an accepted conference paper to the *IEEE International Instrument and Measurement Technology Conference 2018* [145]. However, there is much room to explore the best implementations and uses of GVD connectivity. This includes exploring more deeply the relationship between the stable connectivity estimate and the node function which would hopefully help define ideal, rather than simply recommended, node functions to use for a given connectivity estimate. I am also keen to utilise the GVD connectivity to the important problem of structure-function relationships in neuroimaging, with which I believe there is much scope for interesting analyses by treating the stable connectivity as structural connectivity and the GVD node function as functional connectivity. With such advancements these tools, with high sensitivity to correlated sources, could be of particular use in detecting the source of epileptic seizures in which currently 1/3 of all surgery operations are unsuccessful in removing the correct sources [154].

The methods developed provide interesting new insights into VSTM binding function of two important forms of AD. Pursuing further research in this area is of high interest. This includes gaining access to datasets from larger cohorts to get more powerful evidence of disease abnormalities; applying to other modalities such as fMRI and MRI structural connectivity to get more spatially refined insights into how degree hierarchical topology is linked in structure-function of disease; and employing more powerful classification techniques using machine learning (e.g. from support vector machines to convolutional neural networks however the scale dictates) to aid in detection of disease and move towards practical biomarkers over the general population.

7.4 Conclusions

Overall, the developments in this thesis open up new ways to consider graph-based analysis of data both topologically, using methods to understand and exploit the hierarchical complexity of EEG functional connectivity, and temporally, using a novel framework for graph-based analysis of multivariate signals to explore instantaneous connectivity patterns in a robust way. The

methods in both of these domains show promise in detecting the critical societal problem of AD in both early- and late-onset forms and I expect the impact of these methods to extend to other brain pathologies, as well as other problems which benefit from complex network modelling. I hope the style of enquiry elicited will inspire others to challenge established or growing ideas and perspectives in the science community to promote creativity and lateral thinking and ward off the dulling effects of institutionalisation.

Appendix A: List of publications

1. **K. Smith**, H. Azami, M.A. Parra, J.M. Starr, J. Escudero, "Cluster-span threshold: an unbiased threshold for binarising weighted complete networks in functional connectivity analysis", *IEEE Proceedings of the EMBC2015*, 2840-2843, 2015.
2. **K. Smith**, H. Azami, J. Escudero, M.A. Parra, J.M. Starr, "Comparison of network analysis approaches on EEG connectivity in Beta during visual short-term memory tasks", *IEEE Proceedings of the EMBC2015*, 2207-2210, 2015.
3. **K. Smith**, B. Ricaud, N. Shahid, M.A. Parra, J.M. Starr, J. Escudero, P. Vandergheynst, "Distinguishing EEG activity in visual short-term memory tasks using modular Dirichlet energy", *OHBM 2016 Annual Meeting*, abstract 2086, 2016.
4. **K. Smith**, D. Abásolo, J. Escudero, "A comparison of the cluster-span threshold and the union of shortest paths as objective thresholds of EEG functional connectivity networks from Beta activity in Alzheimer's disease", *IEEE Proceedings of the EMBC2016*, 2826-2829, 2016.
5. J. Escudero, **K. Smith**, H. Azami, D. Abásolo, "Inspection of short-time resting-state electroencephalogram functional networks in Alzheimer's disease", *IEEE Proceedings of the EMBC2016*, 2810-2813, 2016.
6. **K. Smith**, J. Escudero, "Weighted Complex Modular Hierarchy Model", *The 5th International Workshop on Complex Networks & Their Applications*, Book of abstracts: 18-20, 2016.
7. **K. Smith**, J. Escudero, "The complex hierarchical topology of EEG functional connectivity", *Journal of Neuroscience Methods*, **276**: 1-12, 2017.
8. **K. Smith**, B. Ricaud, N. Shahid, S. Rhodes, J. Starr, A. Ibanez, M.A. Parra M.A, J. Escudero, P. Vandergheynst, "Locating temporal functional dynamics in visual short-term memory tasks using modular Dirichlet energy", *Scientific Reports*, **7**: 42013, 2017.
9. **K. Smith**, D. Abásolo, J. Escudero, "Accounting for the complex hierarchical topology

of EEG phase-based functional connectivity in network binarisation", *PLOS ONE*, **12(10)**: e0186164, 2017.

10. **K. Smith**, L. Spyrou, J. Escudero, "Graph-variate signal analysis", pre-print: `https://arxiv.org/abs/1703.06687`, under revision, 2018.

11. Y. Shen, C. Tan, **K. Smith**, J. Escudero, F. Dong, Gas-water two-phase flow pattern recognition based on ERT and ultrasound Doppler, *accepted to IEEE I2MTC2018*, 2018.

Bibliography

[1] K. Smith and J. Escudero, "The complex hierarchical topology of EEG functional connectivity," *Journal of Neuroscience Methods*, vol. 276, pp. 1–12, 2017.

[2] K. Smith, D. Abásolo, and J. Escudero, "Accounting for the complex hierarchical topology of EEG phase-based functional connectivity in network binarisation," *PLOS ONE*, vol. 12, no. 10, p. e0186164, 2017.

[3] K. Smith, L. Spyrou, and J. Escudero, "Graph-variate signal analysis," 2017. [Online]. Available: https://arxiv.org/abs/1703.06687

[4] K. Smith, B. Ricaud, N. Shahid, S. Rhodes, J. Starr, A. Ibanez, M. Parra, J. Escudero, and P. Vandergheynst, "Locating temporal functional dynamics in visual short-term memory tasks using modular Dirichlet energy," *Scientific Reports*, vol. 7, p. 42013, 2017.

[5] R. Shields, "Cultural topology: the seven bridges of Königsberg 1736," *Theory Culture and Society*, vol. 29, no. 4-5, pp. 43–57, 2012.

[6] M. Newman, *Networks*. Oxford, UK: Oxford University Press, 2010.

[7] K. Appel and W. Haken, "Every planar map is four colorable. Part 1: Discharging," *Illinois Journal of Mathematics*, vol. 21, no. 3, pp. 429–490, 1977.

[8] R. Read and D. Corneil, "The graph isomorphism disease," *Journal of Graph Theory*, vol. 1, no. 4, pp. 339–363, 1977.

[9] F. Vega-Redondo, *Complex Social Networks*. Cambridge University Press, 2007.

[10] A. Broder, R. Kumar, F. Maghoui, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, "Graph structure in the web," *Computer Networks*, vol. 33, pp. 309–320, 2000.

[11] M. Girvan and M. Newman, "Community structure in social and biological networks," *PNAS*, vol. 99, no. 12, pp. 7821–7826, 2001.

[12] W. Souma, Y. Fujiwara, and H. Aoyama, "Complex networks and economics," *Physica A*, vol. 324, no. 1-2, pp. 396–401, 2003.

[13] M. Kubat, *An introduction to machine learning*. Springer, 2017, ch. Artificial neural networks, pp. 91–111.

[14] R. Cancho and R. Solé, "The small world of human language," *Proceedings of the Royal Society B*, vol. 268, no. 1482, p. doi:10.1098/rspb.2001.1800, 2001.

[15] *Illustration of a neuron*, Std. [Online]. Available: https://commons.wikimedia.org/wiki/File:Neuron_-_annotated.svg

[16] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature Neuroscience Review*, vol. 10, pp. 186–198, 2009.

[17] B. Tijms, A. Wink, W. De Haan, W. Van der Flier, C. Stam, P. Scheltens, and F. Barkhof, "Alzheimer's disease: connecting findings from graph theoretical studies of brain networks," *Neurobiology of Ageing*, vol. 34, pp. 2023–2036, 2013.

[18] B. Bernhardt, S. Hong, A. Bernasconi, and N. Bernasconi, "Imaging structural and functional brain networks in temoral lobe epilepsy," *Frontiers in Human Neuroscience*, vol. 7, p. 624, 2013.

[19] V. Calhoun, T. Eichele, and G. Pearlson, "Functional brain networks in schizophrenia: a review," *Frontiers in Human Neuroscience*, vol. 3, p. 17, 2009.

[20] "Report of the committee on methods of clinical examination in electroencephalography: 1957," *Electroencephalography and Clinical Neurophysiology*, vol. 10, no. 2, pp. 370–375, 1958.

[21] F. Da Silva, "EEG and MEG: relevance to neuroscience," *Neuron*, vol. 80, no. 5, pp. 1112–1128, 2013.

[22] J. Dauwels, F. Vialatte, and A. Cichocki, *Advances in cognitive neurodynamics (II)*. Springer Netherlands, 2011, ch. On the early diagnosis of Alzheimer's disease from EEG signals: a mini-review, pp. 709–716.

[23] J. Snaedal, G. Johannesson, T. Gudmundsson, S. Gudmundsson, T. Pajdak, and K. Johnsen, "The use of EEG in Alzheimer's disease, with and witwith scopolamine- a pilot study," *Clinical Neurophysiology*, vol. 121, no. 6, pp. 836–841, 2010.

[24] E. Van Straaten, P. Scheltens, A. Gouw, and C. Stam, "Eyes-closed task-free electroencephalography in clinical trials for Alzheimer's disease: an emerging method based upon brain dynamics," *Alzheimer's Research & Therapy*, vol. 6, no. 9, p. 86, 2014.

[25] S. Van den Broek, F. Reinders, M. Donderwinkel, and M. Peters, "Volume conduction effects in EEG and MEG," *Electroencephalography and Clinical Neurophysiology*, vol. 106, no. 6, pp. 522–534, 1998.

[26] F. De Vico Fallani, J. Richiardi, M. Chavez, and S. Achard, "Graph analysis of functional brain networks: practical issues in translational neuroscience," *Philosophical Transactions of the Royal Society B*, vol. 369, no. 1653, p. 20130521, 2014.

[27] D. Papo, J. Buldu, S. Boccaletti, and E. Bullmore, "Complex network theory and the brain," *Philosophical Transactions of the Royal Society B*, vol. 369, p. 20130520, 2014.

[28] T. Rappaport, *Wireless communication: principles and practices*. IEEE Press, 1996.

[29] J. Dauwels, F. Viallate, T. Musha, and A. Cichocki, "A comparative study of synchrony measures for the early diagnosis of Alzheimer's disease based on EEG," *NeuroImage*, vol. 49, no. 1, pp. 668–693, 2010.

[30] E. Van Diessen, T. Numan, A. Van Dellen, A. Van der Kooi, M. Boersma, D. Hofman, R. Van Lutterverld, B. Van Dijk, E. Van Straaten, A. Hillebrand, and C. Stam, "Opportunities and methodological challengs in EEG and MEG resting state functional brain network research," *Clinical Neurophysiology*, vol. 126, no. 8, pp. 1468–1481, 2015.

[31] C. Stam, G. Nolte, and A. Daffertshofer, "Phase-lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources," *Human Brain Mapping*, vol. 28, no. 11, pp. 1178–1193, 2007.

[32] C. Stam, W. de Haan, A. Daffertshofer, B. Jones, I. Manshanden, A. Van Cappellen van Walsum, T. Montez, J. Verbunt, J. de Munck, B. van Dijk, H. Berendse, and P. Scheltens, "Graph theoretical analysis of magnetoencephalographic functional commectivity in Alzheimer's disease," *Brain*, vol. 132, no. 1, pp. 213–224, 2009.

[33] M. Engels, S. C.J., W. Van der Flier, P. Scheltens, H. de Waal, and E. van Straaten, "Declining functional connectivity and changing hub locations in Alzheimer's disease: an EEG study," *BCM Neurology*, vol. 15, no. 145, pp. doi:10.1186/s12 883–015–0400–7, 2015.

[34] M. Yu, A. Gouw, A. Hillebrand, B. Tijms, S. C.J., E. van Straaten, and Y. Pijnenburg, "Different functional connectivity and network topology in behavioural variant of frontotemporal dementia and Alzheimer's disease: an EEG study," *Neurobiology of Aging*, vol. 42, pp. 150–162, 2016.

[35] M. Yu, M. Engels, A. Hillebrand, E. van Straaten, A. Gouw, C. Teunissen, V. der Flier W.M., P. Scheltens, and C. Stam, "Selective impairment of hippocampus and posterior hub areas in Alzheimer's disease: an MEG based multiplex network study," *Brain*, vol. 140, no. 5, pp. 1466–1485, 2017.

[36] M. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Phsical Review E*, vol. 69, no. 2, p. 026113, 2004.

[37] C. Stam, B. Jones, G. Nolte, M. Breakspear, and P. Scheltens, "Small-world neworks and functional connectivity in Alzheimer's disease," *Cerebral Cortex*, vol. 17, pp. 92–99, 2007.

[38] S. Achard and E. Bullmore, "Efficiency and cost of economical brain functional networks," *PLOS ONE*, vol. 3, no. 2, p. e17, 2007.

[39] M. Rubinov and O. Sporns, "Weight-conserving characterization of complex functional brain networks," *NeuroImage*, vol. 56, no. 4, pp. 2068–2079, 2011.

[40] B. Van Wijk, C. Stam, and A. Daffertshofer, "Comparing brain networks of different size and connectivity density using graph theory," *PLOS ONE*, vol. 5, no. 10, p. e13701, 2010.

[41] C. Stam, "Modern network science of neurological disorders," *Nature Neuroscience Review*, vol. 15, pp. 683–695, 2014.

[42] K. Smith, D. Abasolo, and J. Escudero, "A comparison of the cluster-span threshold and the union of shortest paths as objective thresholds of EEG functional connectivity networks from Beta activity in Alzheimer's disease," in *IEEE Proceedings of the EMBC2016*, 2016, pp. 2826–2829.

[43] C. Stam, P. Tewarie, E. Van Dellen, E. Van Straaten, A. Hillebrand, and P. Van Mieghem, "The trees and the forest: characterization of complex brain networks with minimum spanning trees," *International journal of psychophysiology*, vol. 92, pp. 129–138, 2014.

[44] P. Tewarie, E. van Dellen, A. Hillebrand, and C. Stam, "The minimum spanning tree: an unbiased method for brain network analysis," *NeuroImage*, vol. 104, pp. 177–188, 2015.

[45] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: uses and interpretations," *NeuroImage*, vol. 52, pp. 1059–1069, 2010.

[46] J. Kruskal, "On the shortest spanning subtree of a graph and the traveling salesman problem," *American Mathematical Society*, vol. 7, no. 1, pp. 48–50, 1956.

[47] J. Meier, P. Tewarie, and P. van Mieghem, "The union of shortest path trees of functional brain networks," *Brain Connectivity*, vol. 5, no. 9, pp. 575–581, 2015.

[48] E. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Mathematik*, vol. 1, no. 1, pp. 269–271, 1959.

[49] K. Smith, H. Azami, M. Parra, J. Starr, and J. Escudero, "Cluster-span threshold: an unbiased threshold for binarising weighted complete networks in functional connectivity analysis," in *IEEE Proceedings of the EMBC2015*, 2015, pp. 2840–2843.

[50] F. De Vico Fallani, V. Latora, and M. Chavez, "A topological criterion for filtering information in complex brain networks," *PLOS Computational Biology*, vol. 13, no. 1, p. e1005305, 2017.

[51] E. Bullmore and O. Sporns, "The economy of brain network organisation," *Nature Reviews Neuroscience*, vol. 13, pp. 336–349, 2012.

[52] D. Watts and S. Strogatz, "Collective dynamics of small-world networks," *Nature*, vol. 393, pp. 440–442, 1998.

[53] O. Sporns, "Small-world connectivity, motif composition and complexity of fractal neuronal connections," *BioSystems*, vol. 85, pp. 55–64, 2006.

[54] D. Bassett and E. Bullmore, "Small-world brain networks," *Neuroscientist*, vol. 12, no. 6, pp. 512–523, 2006.

[55] S. C.J. and E. Van Straaten, "The organization of physiological brain networks," *Clinical Neurophysiology*, vol. 123, no. 6, pp. 1067–1087, 2012.

[56] A. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, 1999.

[57] V. Eguiluz, D. Chialvo, G. Cecchi, M. Baliki, and A. Apkarian, "Scale free brain functional networks," *Physical Review Letters*, vol. 94, p. 018102, 2005.

[58] M. van den Heuvel, C. Stam, M. Boersma, and H. Hulshoff Pol, "Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain," *NeuroImage*, vol. 43, no. 3, pp. 528–539, 2008.

[59] S. Achard, R. Salvador, B. Whitcher, J. Suckling, and E. Bullmore, "A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs," *Journal of Neuroscience*, vol. 26, no. 1, pp. 63–72, 2006.

[60] M. Van Den Heuvel and O. Sporns, "Network hubs in the human brain," *Trends in Cognitive Sciences*, vol. 17, no. 12, pp. 683–696, 2013.

[61] M. Newman, "Modular and community structure in networks," *Physical Review E*, vol. 23, pp. 8577–8582, 2006.

[62] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 10, p. P10008, 2008.

[63] M. Chavez, M. Valencia, V. Navarro, V. Latora, and J. Matinerie, "Functional modularity of background activities in normal and epileptic brain networks," *Physical Review Letters*, vol. 104, no. 11, p. 118701, 2010.

[64] D. Meunier, R. Lambiotte, and E. Bullmore, "Modular and hierarchical modular organisation of brain networks," *Frontiers in Neuroscience*, vol. 4, p. 200, 2010.

[65] R. Solé and S. Valverde, *Complex Networks*, ser. Lecture Notes in Physics. Springer, 2004, vol. 650, ch. Information theory of complex networks: on evolution and architectural constraints, pp. 189–207.

[66] O. Sporns, *Networks of the brain*. Masachusettes, USA: MIT Press, 2010.

[67] M. Kaiser, C. Hilgetag, and R. Kötter, "Hierarchy and dynamics of neural networks," *Frontiers in Neuroinformatics*, vol. 4, p. 112, 2010.

[68] C. Shannon, "A mathematical theory of communication," *Bell Systems Technical Journal*, vol. 27, pp. 623–656, 1948.

[69] V. Calhoun, R. Miller, G. Pearlson, and T. Adali, "The chronnectome: time-varying connectivity networks as the next frontier in fMRI data discovery," *Neuron*, vol. 84, no. 2, pp. 262–274, 2014.

[70] M. Parra, S. Abrahams, K. Fabi, R. Logie, S. Luzzi, and S. Della Salla, "Short-term memory binding defecits in Alzheimer's disease," *Brain*, vol. 132, no. 4, pp. 1057–1066, 2009.

[71] P. Holme and J. Saramäki, "Temporal networks," *Physics Reports*, vol. 519, no. 3, pp. 97–125, 2012.

[72] M. De Domenico, A. Solé-Ribalta, E. Cozzo, M. Kivelä, Y. Moreno, M. Porter, S. Gómez, and A. Arenas, "Mathematical formulation of multi-layer networks," *Physical Review X*, vol. 3, p. 041022, 2013.

[73] K. Doron, D. Bassett, and M. Gazzaniga, "Dynamic network structure of interhemispheric coordination," *PNAS*, vol. 109, no. 46, pp. 18 661–18 668, 2012.

[74] N. Leonardi and D. Van de Ville, "On spurious and real fluctuations of dynamic functional connectivity during rest," *NeuroImage*, vol. 104, pp. 430–436, 2015.

[75] U. Braun, A. Schäfer, H. Walter, S. Erk, N. Romanczuk-Seiferth, L. Haddad, J. Schweiger, O. Grimm, A. Heinz, H. Tost, A. Meyer-Lindenberg, and D. Bassett, "Dynamic reconfiguration of frontal brain networks during executive cognition in humans," *PNAS*, vol. 112, no. 37, pp. 11 678–11 683, 2015.

[76] U. Braun, A. Schäfer, D. Bassett, F. Rausch, J. Schweiger, E. Bilek, S. Erk, N. Romanczuk-Seiferth, O. Grimm, L. Geiger, L. Haddad, K. Otto, S. Mohnke, A. Heinz, M. Zink, H. Walter, A. Schwarz, A. Meyer-Lindenberg, and H. Tost, "Dynamic brain network reconfiguration as a potential schizophrenia genetic risk mechanism modulated by nmda receptor function," *PNAS*, vol. 113, no. 44, pp. 12 568–12 573, 2016.

[77] E. Glerean, J. Salmi, J. Lahnkoski, I. Jäskel"ainen, and M. Sams, "Functional magnetic resonance imaging phase synchronization as a measure of dynamic functional connectivity," *Brain Connectivity*, vol. 20122, no. 2, pp. 91–101, 2011.

[78] D. Shuman, S. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs," *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 83–98, 2013.

[79] A. Sandryhaila and J. Moura, "Discrete signal processing on graphs," *IEEE Transactions on Signal Processing*, vol. 61, no. 7, pp. 1644–1656, 2013.

[80] ——, "Big data analysis with signal processing on graphs: representation and processing of massive data sets with irregular structure," *IEEE Signal Processing Magazine*, vol. 31, no. 5, pp. 80–90, 2014.

[81] L. Riu, H. Nejati, and N. Cheung, "Dimensionality reduction of brain imaging data using graph signal processing," in *IEEE ICIP*, 2016, pp. 1329–1333.

[82] A. Loukas and D. Foucard, "Frequency analysis of temporal graph signals," in *IEEE Global Conference on Signal and Information Processing*, 2016, pp. 346–350.

[83] F. Grassi, A. Loukas, N. Perraudin, and B. Ricaud, "A time-vertex signal processing framework," 2017. [Online]. Available: arXiv:1705.02307

[84] M. Meyer-Luehmann, T. Spirse-Jones, C. Prada, M. Garcia-Alloza, A. de Calignon, A. Rozkalne, J. Koenigsknecht-Talboo, D. Holtzman, B. Bacskai, and B. Hyman, "Rapid appearance and local toxicity of amyloid-β plaques in a mouse model of Alzheimer's disease," *Nature*, vol. 451, pp. 720–724, 2008.

[85] A. Delacourte and A. Defossez, "Alzheimer's disease: Tau proteins, the promoting factors of microtubule assembly, are major components of paired helical filaments," *Journal of the Neurological Sciences*, vol. 76, no. 2-3, pp. 172–186, 1986.

[86] A. Goate, M. Chartier-Harlin, M. Mullan, J. Brown, F. Crawford, L. Fidani, L. Giuffra, A. Haynes, N. Irving, L. James, R. Mant, P. Newton, K. Rooke, P. Roques, C. Talbot, M. Pericak-Vance, A. Roses, R. Williamson, M. Rossor, M. Owen, and J. Hardy, "Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease," *Nature*, vol. 349, pp. 704–706, 1991.

[87] P. Vemuri and C. Jack Jr, "Role of structural mri in Alzheimer's disease," *Alzheimer's Research & Therapy*, vol. 2, p. 23, 2010.

[88] D. Selkoe, "Alzheimer's disease is a synaptic failure," *Science*, vol. 298, no. 5594, pp. 789–791, 2002.

[89] W. De Haan, Y. Pijnenburg, R. Strijers, Y. van der Made, W. Van der Flier, P. Scheltens, and S. C.J., "Functional neural network analysis in frontotemporal dementia and Alzheimer's disease using EEG and graph theory," *BMC Neuroscience*, vol. 10, p. 101, 2009.

[90] R. Buckner, J. Sepulcre, T. Talukdar, F. Krienen, H. Liu, T. Hedden, J. Andrews-Hanna, R. Sperling, and K. Johnson, "Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease," *Journal of Neuroscience*, vol. 29, no. 6, pp. 1860–1873, 2009.

[91] X. Miao, X. Wu, R. Li, C. K, and L. Yao, "Altered connectivity patterns of hubs in default-mode network with Alzheimer's disease: an Granger causality modelling approach," *PLOS ONE*, vol. 6, no. 10, p. e25546, 2011.

[92] W. De Haan, K. Mott, E. Van Straaten, P. Scheltens, and S. C.J., "Activity dependent degeneration explains hub vulnerability in Alzheimer's disease," *PLOS Computational Biology*, vol. 8, no. 8, p. e1002582, 2012.

[93] Y. Liu, C. Yu, X. Zhang, J. Liu, Y. Duan, A. Alexander-Bloch, B. Liu, T. Jiang, and E. Bullmore, "Impaired long distance functional connectivity and weighted network architecture in Alzheimer's disease," *Cerebral Cortex*, vol. 24, no. 6, pp. 1422–1435, 2014.

[94] Z. Dai, C. Yan, K. Li, Z. Wang, J. Wang, M. Cao, Q. Lin, N. Shu, and a. Xia, M., "Idendity and mapping connectivity patterns of brain network hubs in Alzheimer's disease," *Cerebral Cortex*, vol. 25, no. 10, pp. 3723–3742, 2015.

[95] M. Wheeler and A. Treisman, "Binding in short-term visual memory," *Journal of Experimental Psychology*, vol. 131, no. 1, pp. 48–64, 2002.

[96] M. Parra, S. Abrahams, R. Logie, and S. Della Salla, "Visual short-term memory binding in Alzheimer's disease and depression," *Journal of Neurology*, vol. 257, no. 7, pp. 1160–1169, 2010.

[97] M. Parra, S. Abrahams, R. Logie, L. Méndez, F. Lopera, and S. Della Salla, "Visual short-term memory binding defecits in familial Alzheimer's disease," *Brain*, vol. 133, no. 9, pp. 2702–2713, 2010.

[98] M. Costa, A. Goldberger, and C. Peng, "Multiscale entropy analysis of biological signals," *Physical Review E*, vol. 71, p. 021906, 2005.

[99] G. Tononi, O. Sporns, and G. Edelman, "A measure for brain complexity: relating functional segregation and integration in the nervous system," *PNAS*, vol. 91, no. 11, pp. 5033–5037, 1994.

[100] C. Li, H. Wang, W. de Haan, C. J. Stam, and P. V. Mieghem, "The correlation of metrics in complex networks with applications in functional brain networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2011, no. 11, p. P11018, 2011.

[101] A. Barabási, R. Albert, and H. Jeong, "Mean-field theory for scale-free random networks," *Physica A*, vol. 272, no. 1-2, pp. 173–187, 1999.

[102] T. Snijders, "The degree variance: an index of graph heterogeneity," *Social Networks*, vol. 3, no. 3, pp. 163–174, 1981.

[103] J. McAuley, L. Costa, and T. Caetano, "The rich-club phenomena across complex network hierarchies," *Applied Physics Letters*, vol. 91, p. 084103, 2007.

[104] P. Erdös and A. Rényi, "On random graphs," *Pubilcationes Mathematicae Debrecen*, vol. 6, pp. 290–297, 1959.

[105] P. Bonacich and P. Lloyd, "Eigengraphs-like measures of centrality for asymmetric relations," *Social Networks*, vol. 23, no. 3, pp. 191–201, 2001.

[106] M. Molloy and B. Reed, "A critical point for random graphs with a given degree sequence," *Random Structures & Algorithms*, vol. 6, no. 2-3, pp. 161–180, 1995.

[107] S. Milgram, "The small world problem," *Psychology Today*, vol. 1, no. 1, pp. 61–67, 1967.

[108] J. Petersen, "Die theorie der regulären graphs," *Acta Mathematica*, vol. 15, pp. 193–220, 1891.

[109] B. Ábrego, S. Fernández-Merchant, M. Neubauer, and W. Watkins, "Sum of squares of degree in a graph," *Journal of Inequalities in Pure and Applied Mathematics*, vol. 10, no. 3, pp. 92–99, 2009.

[110] M. Lynall, D. Bassett, R. Kerwin, P. McKenna, M. Kitzbichler, U. Muller, and E. Bullmore, "Functional connectivity in brain networks in schizophrenia," *Journal of Neuroscience*, vol. 30, no. 28, pp. 9477–9487, 2010.

[111] C. Ginestet, T. Nichols, E. Bullmore, and A. Simmons, "Brain nework analysis: separating cost from topology using cost-integration," *PLOS ONE*, vol. 6, p. e21570, 2011.

[112] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. Ivanov, R. Mark, J. Mietus, G. Moody, C. Peng, and H. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[113] G. Schalk, D. McFarland, T. Hinterberger, N. Birbaumer, and J. Wolpaw, "Bci2000: a general purpose brain-computer interface (BCI) system," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1034–1043, 2004.

[114] R. Oostenveld, P. Fries, E. Maris, and J. Schoffelen, "Fieldtrip: open source software for advanced analysis of MEG, EEG, and invasive electrophsiological data," *Computational Intelligence and Neuroscience*, vol. 2011, p. 156869, 2011.

[115] M. Vinck, R. Oostenveld, M. van Wingerden, F. Battaglia, and C. Pennartz, "An improved index of phase-synchronization for electrophysiological data in the presence of volume conduction, noise and ample-size bias," *NeuroImage*, vol. 55, pp. 1548–1565, 2011.

[116] M. Van den Heuvel and O. Sporns, "Rich-club organisation of the human connectome," *Journal of Neuroscience*, vol. 31, no. 44, pp. 15 775–15 786, 2011.

[117] K. Smith, H. Azami, J. Escudero, M. Parra, and J. Starr, "Comparison of network analysis approaches on EEG connectivity in Beta during visual short-term memory tasks," in *Proceedings of the IEEE EMBC15*, 2015, pp. 2207–2210.

[118] A. Schwarz and J. McGonigle, "Negative edges and soft thrthreshold in complex network analysis of resting state functional connectivity data," *NeuroImage*, vol. 55, no. 3, pp. 1132–1146, 2011.

[119] K. Garrison, D. Scheinost, E. Finn, X. Shen, and R. Constable, "The (in)stability of functional brain network measures across thresholds," *NeuroImage*, vol. 118, pp. 651–661, 2015.

[120] M. Jalili, "Functional brain networks: does the choice of dependency estimator and binarization method matter?" *Scientific Reports*, vol. 6, p. 29780, 2016.

[121] M. Van den Heuvel, S. de Lange, A. Zalesky, C. Seguin, T. Yeo, and R. Schmidt, "Proportional thresholding in resting-state fMRI functional connectivity networks and consequences for patient-control connectome studies: issues and recommendations," *NeuroImage*, vol. 152, pp. 437–449, 2017.

[122] T. Tanizawa, G. Paul, R. Cohen, S. Havlin, and H. Stanley, "Optimization of network robustness to waves of targeted and random attacks," *Physical Review E*, vol. 71, p. 047101, 2005.

[123] C. Dormann, J. Fründ, N. Blüthgen, and B. Gruber, "Indices, graphs and null models: analyzing bipartite ecological networks," *Open Ecology Journal*, vol. 2, pp. 7–24, 2009.

[124] *Neurophysiological Biomarker Toolbox*, Std., 2017. [Online]. Available: www.nbtwiki.net

[125] J. Escudero, D. Abásolo, R. Hornero, P. Espino, and M. López, "Analysis of electroencephalograms in Alzheimer's disease patients with multiscale entropy," *Physiological Measurement*, vol. 27, no. 11, pp. 1091–1106, 2016.

[126] R. Barry, A. Clarke, S. Johnstone, C. Magee, and J. Rushby, "EEG differences between eyes-closed and eyes-open resting conditions," *Clinical Neurophysiology*, vol. 118, no. 12, pp. 2765–2773, 2007.

[127] M. Parra, S. Della Salla, R. Logie, and A. Morcom, "Neural correlates of shape-color binding in visual working memory," *Neuropsychologia*, vol. 52, no. 0, pp. 27–36, 2014.

[128] M. Pietto, M. Parra, N. Trujillo, F. Flores, A. García, J. Bustin, P. Richly, F. Manes, I. n. A. Lopera, F., and S. Baez, "Behavioural and electrophysiological correlates of memory binding defecits in patients at different risk levels for Alzheimer's disease," *Journal of Alzheimer's Disease*, vol. 53, no. 4, pp. 1325–1340, 2016.

[129] M. Folstein, S. Folstein, and P. McHugh, "Mini-mental state: a practical method for grading the cognitive state of patients for the clinician," *Journal of Psychiatric Research*, vol. 12, pp. 189–198, 1975.

[130] M. Kaiser, "Mechanisms of connectome development," *Trends in Cognitive Sciences*, vol. 21, no. 9, pp. 703–717, 2017.

[131] E. Santarnecchi, G. Galli, N. Polizzotto, A. Rossi, and S. Rossi, "Efficiency of weak brain connections support general cognitive functioning," *Human Brain Mapping*, vol. 35, no. 9, pp. 4566–4582, 2014.

[132] K. Joyce, S. Hayasaka, and P. Laurienti, "The human functional brain network demonstrates structual and dynamical resilience to targeted attack," *PLOS Computational Biology*, vol. 9, no. 1, p. e1002885, 2013.

[133] U. Lee, G. Oh, S. Kim, G. Noh, B. Choi, and G. Mashour, "Brain networks maintain a scale-free organization across consciousness, anesthesia, and recovery: evidence for adaptive reconfiguration," *Anesthesiology*, vol. 113, pp. 1081–1091, 2010.

[134] W. Schoen, J. Chang, U. Lee, P. Bob, and G. Mashour, "The temporal organization of functional brain connectivity is abnormal in schizophrenia but does not correlate with symptomatology," *Conscious Cognition*, vol. 20, pp. 1050–1054, 2011.

[135] M. Boersma, D. Smit, D. Boomsma, E. de Geus, H. Delemarre-van de Waal, and C. Stam, "Growing trees in child brains: graph theoretical analysis of electroencephalography-derived minimum spanning tree in 5- and 7-year-old children reflects brain maturation," *Brain Connectivity*, vol. 3, pp. 50–60, 2013.

[136] D. Shuman, C. Wiesmeyr, N. Holighaus, and P. Vanderheynst, "Spectrum-adapted tight graph wavelet and vertex-frequency frames," *IEEE Transactions on Signal Processing*, vol. 63, no. 16, pp. 4223–4235, 2015.

[137] R. Henson, D. Wakeman, V. Litvak, and K. Friston, "A parametric empirical bayesian framework for the EEG/MEG inverse problem: generative models for multi-subject and multi-modal integration," *Frontiers of Human Neuroscience*, vol. 5, p. 76, 2011.

[138] B. Rossion and S. Caharel, "ERP evidence for the speed of face categorization in the human brain: Disentangling the contribution of low-level visual cues from face perception," *Vision Research*, vol. 51, pp. 1297–1311, 2011.

[139] A. Peled, A. Geva, W. Kremen, H. Blankfeld, R. Esfandiarfard, and T. Nordahl, "Functional connectivity and working memory in schizophrenia: an EEG study," *International Journal of Neuroscience*, vol. 106, no. 1-2, pp. 47–61, 2001.

[140] K. Supekar, V. Menon, D. Rubin, M. Musen, and M. Greicius, "Network analysis of intrinsic functional brain connectivity in Alzheimer's disease," *Computational Biology*, vol. 4, p. e1000100, 2008.

[141] H. Zimmer, "Visual and spatial working memory: from boxes to netwoks," *Neuroscience Biobehavioural Reviews*, vol. 32, no. 8, pp. 1373–1395, 2008.

[142] M. Parra, H. Saarimäki, M. Bastin, A. Londoño, L. Pettit, F. Lopera, S. Della Salla, and S. Abrahams, "Memory binding and white matter integrity in familial Alzheimer's disease," *Brain*, vol. 138, no. 5, pp. 1355–1369, 2015.

[143] D. Yekuteli, "Hierarchical false discovery rate-controlling methodology," *Journal of the American Statistical Association*, vol. 103, no. 481, pp. 309–316, 2008.

[144] A. Griffa, B. Ricaud, K. Benzi, X. Bresson, A. Daducci, P. Vandergheynst, J. Thiran, and P. Hagmann, "Transient networks of spatio-temporal connectivity map communication pathways in brain functional systems," *NeuroImage*, vol. 155, pp. 490–502, 2017.

[145] Y. Shen, C. Tan, K. Smith, J. Escudero, and F. Dong, "Gas-water two-phase flow pattern recognition based on ERT and ultrasound doppler," in *accepted to 2018 IEEE I2MTC*. IEEE, 2018.

[146] L. Minati, T. Edginton, M. Bruzzone, and G. Giaccone, "Reviews: current concepts in Alzheimer's disease: a multidisciplinary review," *American Journal of Alzheimer's Disease and Other Dementias*, vol. 24, no. 2, pp. 95–121, 2009.

[147] A. Ibáñez and M. Parra, "Mapping memory binding onto the connectome's temporal dynamics: toward a combined biomarker for Alzheimer's disease," *Frontiers in Human Neuroscience*, vol. 8, p. 237, 2014.

[148] X. Delbeuck, F. Collette, and M. Van deVan der, "Is Alzheimer's disease a disconnection syndrome? evidence from a crossmodal audio-visual illusory experiment," *Neuropsychologia*, vol. 45, no. 14, pp. 3315–3323, 2007.

[149] R. Petersen, "Mild cognitive impairment as diagnostic entity," *Journal of International Medecine*, vol. 256, no. 3, pp. 183–194, 2004.

[150] B. Winblad, K. Palmer, M. Kivipelto, V. Jelic, L. Faratiglioni, and e. a. Wahlund, L-O, "Mild cognitive impairment- beyond controversies, ttoward a consensus: report of the international working group on mild cognitive impairment," *Journal of International Medecine*, vol. 256, no. 3, pp. 240–246, 2004.

[151] J. Mitchell, R. Arnold, K. Dawson, P. Nestor, and J. Hodges, "Outcome in subgroups of mild cognitive impairment (MCI) is highly predictable using a simple algorithm," *Journal of Neurology*, vol. 256, no. 9, pp. 1500–1509, 2009.

[152] M. Parra, E. Mikulan, N. Trujillo, S. Della Sala, F. Lopera, F. Manes, J. Starr, and A. Ibáñez, "Brain information sharing during visual short-term memory binding yields a memory biomarker for familial Alzheimer's disease," *Current Alzheimer Research*, vol. 14, no. 12, pp. 1335–1347, 2017.

[153] F. Schweitzer, G. Fagiolo, D. Sornette, F. Vega-Redondo, A. Vespignani, and D. White, "Economic networks: The new challenges," *Science*, vol. 325, no. 5939, pp. 422–425, 2009.

[154] *Epilepsy surgery information*, Epilepsy Society Std., 2017. [Online]. Available: https://www.epilepsysociety.org.uk/epilepsy-surgery#.Wuwk2I4VthE