

ASSESSMENT OF FUNCTIONAL CONNECTIVITY IN THE HUMAN BRAIN:
MULTIVARIATE AND GRAPH SIGNAL PROCESSING METHODS

By

Marisel Villafañe-Delgado

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Electrical Engineering — Doctor of Philosophy

2017

ABSTRACT

ASSESSMENT OF FUNCTIONAL CONNECTIVITY IN THE HUMAN BRAIN: MULTIVARIATE AND GRAPH SIGNAL PROCESSING METHODS

By

Marisel Villafañe-Delgado

Advances in neurophysiological recording have provided a noninvasive way of inferring cognitive processes. Recent studies have shown that cognition relies on the functional integration or connectivity of segregated specialized regions in the brain. Functional connectivity quantifies the statistical relationships among different regions in the brain. However, current functional connectivity measures have certain limitations in the quantification of global integration and characterization of network structure. These limitations include the bivariate nature of most functional connectivity measures, the computational complexity of multivariate measures, and graph theoretic measures that are not robust to network size and degree distribution. Therefore, there is a need of computationally efficient and novel measures that can quantify the functional integration across brain regions and characterize the structure of these networks.

This thesis makes contributions in three different areas for the assessment of multivariate functional connectivity. First, we present a novel multivariate phase synchrony measure for quantifying the common functional connectivity within different brain regions. This measure overcomes the drawbacks of bivariate functional connectivity measures and provides insights into the mechanisms of cognitive control not accountable by bivariate measures. Following the assessment of functional connectivity from a graph theoretic perspective, we propose a graph to signal transformation for both binary and weighted networks. This provides the means for characterizing the network structure and quantifying information in the graph by overcoming some drawbacks of traditional graph based measures. Finally, we introduce a new approach to studying dynamic functional connectivity

networks through signals defined over networks. In this area, we define a dynamic graph Fourier transform in which a common subspace is found from the networks over time based on the tensor decomposition of the graph Laplacian over time.

Copyright by
MARISEL VILLAFANE-DELGADO
2017

To my parents and NC.

TABLE OF CONTENTS

LIST OF TABLES	viii
LIST OF FIGURES	ix
Chapter 1 Introduction	1
1.1 Functional Connectivity and Cognitive Control	3
1.2 Methods for Quantifying Functional Connectivity	4
1.3 Organization and Contributions of this Thesis	6
1.3.1 Multivariate Phase Synchrony and Hyperdimensional Geometry	7
1.3.2 Graph to Signal Transform for Weighted Graphs	8
1.3.3 Dynamic Graph Fourier Transform	10
Chapter 2 Background	12
2.1 Bivariate Time-Frequency Phase Synchrony	12
2.2 Graph Theory	13
2.3 Cognitive control experiment	15
2.3.1 Participants	15
2.3.2 Experiment	15
2.3.3 EEG Data Acquisition and Pre-processing	15
Chapter 3 Hypertorus Multivariate Phase Synchrony	17
3.1 Introduction	17
3.2 Background	20
3.2.1 S-estimator	20
3.3 Hyperspherical Phase Synchrony	21
3.4 Proposed Solution	24
3.4.1 Hyperspherical Approach	24
3.4.2 Hypertorus Synchrony	28
3.4.3 Computational Complexity	30
3.5 Statistical assessment of HTS	31
3.5.1 Bias	31
3.5.2 Variance	34
3.5.3 Correction of Bias in \widehat{HTS}^2	37
3.6 Results	39
3.6.1 Assessment of robustness to noise of multivariate phase synchrony measures	39
3.6.2 Effect of number of oscillators on the multivariate synchrony measures	40
3.6.3 Kuramoto Model	43
3.6.4 Rössler oscillator	45
3.6.5 Assessment of topographical sensitivity	50
3.7 Conclusions	61

Chapter 4	Graph to Signal Transform Based on the Resistance Distance and its applications to Functional Connectivity Networks	63
4.1	Introduction	63
4.2	Background	69
4.2.1	Graph Entropy Measures	69
4.2.2	Graph to Signal Transform based on Classical Multidimensional Scaling	70
4.3	Graph to Signal Transformation Based on the Resistance Distance Matrix	72
4.3.1	Resistance distance	72
4.3.2	Classical Multidimensional Scaling based on the Resistance Distance	74
4.3.3	Reconstruction of the original graph	75
4.3.4	Perturbation Analysis	77
4.3.5	Illustration of Graph to Signal Transform	80
4.3.5.1	Binary graphs	80
4.3.5.2	Weighted graphs	83
4.3.5.3	Reconstruction of Weighted Networks	85
4.3.5.4	Robustness to network anomalies	86
4.4	Small-world network characterization	88
4.5	Graph Entropy based on the graph to signal transform	90
4.6	Event detection in temporal networks	95
4.6.1	Tensor Decompositions for Temporal Networks	95
4.6.2	Graph to signal transform based event detection	97
4.7	Characterization of Functional Connectivity Networks	99
4.7.1	Assessment of Graph Information Theoretic Measures in Functional Connectivity Networks	102
4.8	Conclusions	103
Chapter 5	Dynamic Graph Fourier Transform	106
5.1	Background	108
5.1.1	Graph Signal Processing	108
5.1.2	Tucker Decomposition	111
5.2	Dynamic Graph Fourier Transform on Temporal Networks	112
5.3	Results	113
5.3.1	Simulations	114
5.3.2	Dynamic Functional Connectivity Networks	117
5.4	Conclusions	126
Chapter 6	Conclusions and Future Work	127
	BIBLIOGRAPHY	130

LIST OF TABLES

Table 3.1:	Multivariate synchrony (mean \pm st.dev.) in networks of Rössler oscillators.	49
Table 3.2:	Multivariate synchrony (mean \pm st.dev.) for different number of oscillators containing two subnetworks of three oscillators.	50
Table 3.3:	Multivariate synchrony (mean \pm st.dev.) for a network consisting of 12 oscillators for different number of subnetworks composed of three oscillators.	50
Table 3.4:	Statistical significance of error-correct responses obtained from PLV and HTS.	54
Table 4.1:	Reconstruction errors	85
Table 4.2:	Estimated small-world parameters.	101
Table 4.3:	Small-world measure.	102
Table 4.4:	Graph entropy from cognitive control FCNs.	103
Table 4.5:	Correlations between graph entropy and behavioral measures from cognitive control.	103
Table 5.1:	Performance of the proposed \mathbf{L}_T and \mathbf{L}_A	116
Table 5.2:	MSE of the smoothness of $\mathbf{f}_1^{(t)}$ from \mathbf{L}_T and \mathbf{L}_A	116
Table 5.3:	MSE of the smoothness of $\mathbf{f}_2^{(t)}$ from \mathbf{L}_T and \mathbf{L}_A	117

LIST OF FIGURES

Figure 3.1:	\widehat{HTS} (solid lines) and true HTS (dashed lines) for synchrony values of 0, 0.20, 0.40, 0.60, 0.80 and 0.99. $M = 4$ oscillators were simulated with phases θ_i distributed as $VM(0, \kappa)$	33
Figure 3.2:	(a) Theoretical upper bounds for the variance of HTS; and (b) empirical variance of HTS as a function of sample size in a network of $M = 4$ oscillators for different synchronization levels in the Von Mises distribution in (3.35).	36
Figure 3.3:	\widehat{HTS}^2 (solid lines) and true HTS^2 (dashed lines) for true HTS values of 0, 0.20, 0.40, 0.60, 0.80 and 0.99.	38
Figure 3.4:	Variance of \widehat{HTS}^2 (solid lines) and true HTS^2 (dashed lines) for true HTS values of 0, 0.20, 0.40, 0.60, 0.80 and 0.99.	38
Figure 3.5:	Multivariate synchrony for a network of highly synchronized sinusoidal oscillators.	40
Figure 3.6:	Effect of number of oscillators on S-estimator and HTS for mean synchrony values of 0.1, 0.3, 0.5, and 0.7.	42
Figure 3.7:	Effect of number of oscillators (M) on the eigenvalues for a true multivariate synchrony value of 0.4, a) $M = 4$; b) $M = 8$; c) $M = 12$; and d) $M = 16$	43
Figure 3.8:	Comparison of mean and standard deviation of multivariate synchrony (HTS and S-estimator) within a Kuramoto network with $K_c = 2$	45
Figure 3.9:	Eight Rössler networks.	48
Figure 3.10:	SynAmps2-64 EEG system and network under test.	52
Figure 3.11:	Topographical plots showing the sensitivity of HTS and S-estimator at various SNR levels. a), c) and e) HTS, 20 dB, 10 dB and 0 dB, respectively; b), d) and f) S-estimator, 20 dB, 10 dB and 0 dB, respectively. . . .	53
Figure 3.12:	Error-Correct topographical sensitivity of multivariate synchrony in intervals of 25 ms and theta band [4-8 Hz]. (a), (c), (e) and (g) Error-Correct HTS difference. (b), (d), (f) and (h) Error-Correct S-estimator difference. .	56

Figure 3.13:	Topographical plots of p-values from t-test investigating the difference of multivariate synchrony from HTS between intervals. Note: Black regions correspond to lower p-values.	57
Figure 3.14:	Topographical plots of multivariate synchrony from HTS in the 25-75 ms interval. (a) Error responses; (b) Correct responses; (c) Error-Correct responses.	58
Figure 3.15:	ROC curves for HTS and S-Estimator. Probability of detection is based on the multivariate synchrony among FCz and its neighbors whereas the probability of false alarm is based on the multivariate synchrony around CPz.	59
Figure 3.16:	Correlation coefficient between (a) PES, (b) PEA and error-correct multivariate synchrony difference computed using HTS in the ERN interval 25-75 ms, for each electrode.	60
Figure 3.17:	Topographical distribution of p-values obtained from the correlation coefficient between (a) PES, (b) PEA and error-correct multivariate synchrony difference computed using HTS in the ERN interval 25-75 ms, for each electrode. Black refers to more significant.	61
Figure 4.1:	Signal representation of a ring lattice network composed of $N = 128$ nodes. Top: Resistance distance; (a) $K = 2$; (b) $K = 10$. Bottom: Distance \mathbf{D} ; (c) $K = 2$; (d) $K = 10$	82
Figure 4.2:	Erdős-Rényi network signal representation; (a) and (b), Resistance distance, $p = 0.2$ and $p = 0.5$, respectively; (c) and (d), Distance \mathbf{D} , $p = 0.2$ and $p = 0.5$, respectively.	83
Figure 4.3:	Signals constructed from a weighted stochastic block network with probability of attachment $p = 0.3$ using the resistance distance matrix \mathbf{R} . (a) First three components corresponding to a network with 3 blocks; (b) First four components corresponding to a network with 4 blocks.	84
Figure 4.4:	Signals constructed from a weighted Small-World network consisting of $N = 128$ nodes and average degree $K = 6$. (a) Rewiring probability $p = 0.1$; (b) Rewiring probability $p = 0.7$	85
Figure 4.5:	Error of the magnitude spectrum. (a) and (b) Ring lattice network with average degree K equal to 4, 16, and 64 consisting of $N = 128$ and $N = 256$ nodes, respectively; (c) Stochastic block network, 3 clusters and probability of attachment $p = 0.1$, $p = 0.3$, $p = 0.5$	87

Figure 4.6:	Error of the magnitude spectrum from a ring lattice network (a) and a stochastic block network (b) with one anomalous edge whose weight ranged in the intervals $[0.8, 1.2]$, $[0.6, 1.4]$, $[0.4, 1.6]$, $[0.2, 1.8]$, and $[0, 2]$	88
Figure 4.7:	Estimated probability of rewiring p_r in weighted small-world networks. Weights of the small world structure are uniformly distributed in the interval $[0, 1]$ and noise values are uniformly distributed in (a) $[0, 1]$, (b) $[0, 0.25]$	89
Figure 4.8:	Power spectrum of the first graph signal for: (a) Ring network with $K = 4$; (b) Small world network with $p_r = 0.01$, $p_r = 0.05$, and $p_r = 0.1$; (c) Erdős-Rényi network with $p = 0.1$, $p = 0.3$, and $p = 0.6$; (d) Stochastic Block network with $C_k = 2$, $C_k = 6$, and $C_k = 10$ clusters, $N = 300$ nodes for all networks. The frequency axis limits are adjusted in order to better illustrate the spectrum.	91
Figure 4.9:	Comparison of graph entropy measures for (a) Ring network, $K = 2, 4, 8, 16, 32$ ($I_{f^{V_1}}, \alpha = 0.98$ and $\alpha = 1.03$); (b) Small world network, $K = 4$ and probability of rewiring p_r ranging from 0.0001 to 1 ($I_{f^{V_1}}, \alpha = 0.95$, and $\alpha = 1.05$); (c) Erdős-Rényi network, p from 0.05 to 1 in increments of 0.05 ($I_{f^{V_1}}, \alpha = 0.95$, and $\alpha = 1.1$); (d) Stochastic Block network, $C_k = 3, 5, 7, 9$ ($I_{f^{V_1}}, \alpha = 0.95$, and $\alpha = 1.1$), and $N = 300$ nodes for all networks.	94
Figure 4.10:	Computation of graph divergence between (a) small-world network with $K = 4$ and $p = 0.0001$ and another small-world network with increasing p ; (b) Stochastic Block network with 3 blocks and $p = 0.9$ and another Stochastic Block with 3 blocks and different probability of attachment p	95
Figure 4.11:	Detection of an event consisting of a Small-World network whose probability of attachment p changes from that of the default network ($p = 0.01$) at $t = 31$. ROCs are constructed from the proposed method (blue) and adjacency matrix based method (red) for (a) $p = 0.05$; (b) $p = 0.1$; (c) $p = 0.15$; (d) $p = 0.2$	99
Figure 4.12:	Magnitude Spectrum for each signal obtained through network to signal transformation for (a) Error responses; (b) Correct responses.	101
Figure 5.1:	Tensor unfolding.	112
Figure 5.2:	DGFT of a ring network with $N = 100$ nodes and $K = 4$ over $T = 80$ seconds. The graph signal is composed of different components over time, which are extracted by the proposed method.	115
Figure 5.3:	dGFT of the ERN dFCN over the interval $[-25 \text{ ms}, 125 \text{ ms}]$	118

Figure 5.4:	dGFT of the CRN dFCN over the interval [-25 ms, 125 ms].	118
Figure 5.5:	Topoplots from filtered signals in the low, medium and high graph frequency bands. Interval: [-25 ms, 0 ms].	120
Figure 5.6:	Topoplots from filtered signals in the low, medium and high graph frequency bands. Interval: [0 ms, 25 ms].	121
Figure 5.7:	Topoplots from filtered signals in the low, medium and high graph frequency bands. Interval: [25 ms, 50 ms].	122
Figure 5.8:	Topoplots from filtered signals in the low, medium and high graph frequency bands. Interval: [50 ms, 75 ms].	123
Figure 5.9:	Topoplots from filtered signals in the low, medium and high graph frequency bands. Interval: [75 ms, 100 ms].	124
Figure 5.10:	Topoplots from filtered signals in the low, medium and high graph frequency bands. Interval: [100 ms, 125 ms].	125

Chapter 1

Introduction

Cognition and perception are founded on the coordinated activity of neural populations communicating among different specialized brain regions [1]. Neurons that synchronously oscillate in the low and high frequency provide the fundamental mechanism for information transfer [2], allowing coordinated activity in the normally functioning brain [3], [4], [5]. This neural coordination is spatiotemporally dynamic [6], and the oscillatory synchronization among different regions is dynamically adjusted based on the cognitive task [3]. Furthermore, cognitive dysfunctions such as schizophrenia, epilepsy, autism, Alzheimer's disease, and Parkinson's disease have been related to abnormalities in neuronal synchronization [1], [7].

Brain functionality has been argued to be based on functional segregation and integration [8], [9]. Functional segregation establishes that specialized activity occurs due to segregated neuronal populations within dedicated brain regions [10]. Functional integration, on the other hand, consists of the combination of multiple distributed regions and serves as the basis for coherent cognition and behavior [10].

The development of brain imaging techniques has provided the means to non-invasively infer patterns of neural activity in the human brain. Among those techniques are electroencephalography (EEG) and magnetoencephalography (MEG), which measure the scalp electric and magnetic fields generated by electrical activity of neural assemblies composed of thousands of neurons in

the cortex, respectively [11]. Both techniques provide a high temporal resolution (in the order of milliseconds), but lack good spatial resolution. In addition, these methodologies are sensitive to current sources taking place in different locations of the cortex. Particularly, EEG is more sensitive to secondary currents (volume), whereas MEG is more sensitive to primary current sources [11]. Another popular neuroimaging technique is functional Magnetic Resonance Imaging (fMRI), which measures the changes in blood flow and oxygenation in the brain [12]. It provides an excellent spatial resolution (in the order of few millimeters), at the expense of a slower temporal resolution (in the order of seconds). The advancements in fMRI studies have had a great impact on the study of task-based activation and resting state networks.

Brain connectivity encompasses three categories: anatomical, effective and functional connectivity. Anatomical connectivity refers to the physical interconnections among neurons or neuronal elements and can vary depending on the time-scale that it is observed, being quasi-stationary during short-time scales, whereas it is more dynamic at long-temporal scales, due to plasticity. Functional connectivity has been defined as the statistical dependencies among remote neurophysiological events [13]. It does not make any assumptions regarding the underlying structural connections nor the directionality relationships among the regions being assessed. Functional connectivity, as opposed to structural connectivity, can be dynamic at both short and long temporal scales. Effective connectivity, on the other hand, describes the causal relationships between different regions in the brain [13].

In this thesis, we focus on the assessment of functional connectivity, as it has been shown to contribute to the understanding of neural functions in cognition [14]. Functional connectivity was initially defined as the temporal coherence among different neurons, measured by the cross-correlation of spike trains [15], [16]. Furthermore, anomalies in the functional connections between certain regions are indicative of cognitive dysfunctions, including Alzheimer's Disease [17]

and schizophrenia [18].

1.1 Functional Connectivity and Cognitive Control

Cognitive control is thought to be the foundation of intelligent behavior [18], and it is the mechanism that allows for performance adjustments under activities such as perceptual selection, novel information, and realization of errors [19], [20]. An important question in neuroscience is how the brain adjusts behavior by monitoring the performance under certain activities [21]. It has been hypothesized that cognitive control relies on the information carried by neural signals associated to control, which allows the appropriate selections of actions [22] and has particular influence in psychopathology and self-monitoring [23]. Brain areas involved in cognitive control are the anterior cingulate cortex (ACC), dorsal medial prefrontal cortex (mPFC), and some regions in the parietal lobes [19], [23], [24].

Evoked-related potentials (ERPs) related to cognitive control include the N2, feedback-related negativity (FRN), conflict-related negativity (CRN), and error-related negativity (ERN) [19]. Among these, the ERN is an indicator of cognitive control occurring when the individual performs a behavioral error [25], reaching its maximum (negative) amplitude within 25-75 ms after the error. It has been shown that the ERN exhibits its maximum potential at central and frontal-central regions [25]. Based on dipole modeling, it is suggested that the ERN has a medial frontal generator, potentially the ACC [20], [21], [22]. Hall et al. [22] reported that the ERN response amplitude reflects reduced self-monitoring associated with psychopathology and is linked to higher scores on a self-report measure. Other studies have related a reduction in the amplitude of ERN to acute alcohol intoxication [23] and increases in ERN amplitude are related to the degree of error responses [24]. In addition, the ERN amplitude is attenuated in patients that experience damage in the dorsal ACC

[24].

Various functional connectivity resting state fMRI studies have related abnormalities in the functional relationships among brain regions associated with cognitive control, which includes the dorsal ACC, dorsal PFC and regions of the parietal lobe [25], to attention deficit hyperactivity disorder (ADHD) [26]. In an EEG study, Cavanagh et al. [27] have found increased synchronization between electrodes in the medial prefrontal cortex (mPFC) and the lateral prefrontal cortex (lPFC). Furthermore, their results suggest that neural oscillations between mPFC and lPFC might be the underlying mechanism of functional communication involving networks related to action monitoring and cognitive control. It is suggested that errors indicate the action monitoring network that there is a need for increased cognitive control [27]. Recently, Cavanagh et al. [19] showed that the theta band oscillations in the frontal regions are responsible determining the need of cognitive control and adjustments for a given task [19].

1.2 Methods for Quantifying Functional Connectivity

Methods for quantifying functional connectivity include linear correlation, mutual information, coherence, phase-locking value (PLV) and pairwise phase consistency [28]. The most commonly implemented methods are linear correlation and coherence, which are only sensitive to the linear interactions between time series, and thus cannot account for the nonlinear relationships reflected between electrophysiological time series [28]. Mutual information measures suffer from the estimation of the distribution and associated problems such as the size of histogram bins [28].

In order to quantify both linear and nonlinear relationships in the brain signals, phase synchronization, as defined in the context of two chaotic oscillators, has emerged as an alternative method for the assessment of functional connectivity and is quantified through PLV [29], [30], [31]. PLV

as a measure for functional connectivity was introduced by [32], and it estimates the synchrony between two signals by looking at the circular variance of their phase difference across trials. In comparison to other linear and nonlinear methods, PLV has been shown to be more sensitive to nonlinear effects [33]. In addition, this metric has contributed to the assessment of brain rhythms and their related cognitive processes, for example, alpha, beta, delta and theta in the low-frequency bands and gamma bands in the higher frequencies [28], [29], [32], [34].

Although PLV is a promising measure for quantifying functional connectivity, it is still limited by its bivariate nature. Specifically, it does not provide information regarding the integration across multiple regions in the brain. In addition, functional connectivity results from bivariate measures are difficult to interpret and computationally expensive for systems with a large number of regions. In order to overcome these drawbacks, researchers have proposed multivariate phase synchrony measures [35], [36], [37], [38]. Multivariate phase synchrony aims to quantify the global connectivity among a group of oscillators. Previously proposed multivariate measures include the S-Estimator [35], [36] and hyperspherical phase synchrony (HPS) [37], [38], but those methods present some drawbacks such as being computationally complex and having poor topographical sensitivity, respectively.

On the other hand, graph theory has provided the means for characterizing the functional connections in the brain, which is a complex dynamic system [39]. Functional connectivity networks are constructed by considering the different brain regions or electrodes/sensors as nodes and the relationships between different nodes, quantified by bivariate functional connectivity measures such as PLV and correlation, as edges. In this manner, functional connectivity networks can take advantage of the widely available set of techniques for characterizing complex networks. In terms of brain networks, these measures have been grouped as measures of functional segregation and functional integration [10]. Measures of functional segregation include the clustering coefficient,

transitivity, and modularity [10]. On the other hand, measures of functional integration include the characteristic path length and the global efficiency. By computing measures that characterize network structure, such as the small-world measure and the degree distribution, it has been shown that functional connectivity networks exhibit features of complex networks, including the small-world network [39], [40], and both small-world and scale-free networks [41].

Although graph-theoretic measures have contributed greatly to the advancements in the study of functional connectivity networks, these measures present some drawbacks. Measures employed in the characterization of network structure such as the mean clustering coefficient, the characteristic path length, and the global efficiency may be affected by certain characteristics of the network. Examples include how nodes with low degree affect the clustering coefficient and the dependence of the characteristic path length and the global efficiency of the shortest path between nodes, when networks may rely on other mechanisms than the shortest path for communication.

1.3 Organization and Contributions of this Thesis

In this thesis, we present novel techniques that aim to overcome some of the drawbacks previously mentioned in the quantification of functional connectivity and the assessment of functional connectivity networks. Extending bivariate functional connectivity measures, we introduce a multivariate phase synchrony measure based on hyperdimensional geometry. Along the lines of graph theory, we introduce a graph to signal transform and a dynamic graph Fourier transform as alternative methods in the study of functional connectivity networks.

1.3.1 Multivariate Phase Synchrony and Hyperdimensional Geometry

Since cognitive processes involve the coordinated activity of multiple regions, functional connectivity measures that account for global integration are needed. Multivariate phase synchrony accounts for the global synchronization of a group of oscillators. In addition, it allows for the quantification of the connectivity structure with a single number instead of a matrix of pairwise values, which means that we can assess the global functional connectivity within a neighborhood of a particular region and represent it with a single number. Furthermore, it provides the means for quantifying the integration of large-scale synchronization and functional connectivity.

In Chapter 3, we introduce a novel multivariate phase synchrony measure in order to overcome the drawbacks of bivariate measures such as PLV in the quantification of functional connectivity among multiple brain regions. The proposed measure overcomes drawbacks of current multivariate phase synchrony measures, such as being computationally efficient, robust to noise, and providing excellent topographical sensitivity. This method, referred to as HyperTorus Synchrony (HTS), is based on defining phase differences within a group of oscillators on a flat hyperdimensional torus, which can be considered as an extension of PLV since it is defined as the Cartesian product of unit circles. This novel measure also accounts for the dependency of the multivariate synchrony in the ordering of the phase differences observed in a recently proposed multivariate phase synchrony, HPS [37], [42]. In this chapter, we show that the proposed HTS is equivalent to an extended version of HPS which includes the coordinates of circles with varying radii and thus is independent of the ordering of the phase differences.

This chapter provides an extensive mathematical characterization of the measure as well as the statistical properties of the proposed estimator. It is shown that the proposed measure is more robust to noise, is not affected by the number of oscillators, and possesses an excellent topographical

sensitivity. Simulation results motivate the use of HTS for the quantification of global functional connectivity in brain networks. In addition, in this chapter we assess the global functional connectivity in cognitive control. It is of particular interest to study the difference between the multivariate synchrony of error responses and correct responses during the ERN interval and its topographical distribution. In order to accomplish this, we study the topographical connectivity by computing the multivariate synchrony in the neighborhood of each electrode and assigning this value to the location of the electrode. In this way, it is possible to construct a multivariate synchrony topographical map, which allows to quantify the integration of multiple regions across the brain.

A comparison of HTS to a conventional multivariate synchrony measure, the S-estimator, reveals high error minus correct differences within fronto-lateral and medial-central regions from HTS. In addition, by computing the multivariate synchrony from HTS over different time intervals we identify significant changes in the connectivity within central regions. Furthermore, by summarizing the global synchronization within a region by a single number at each electrode we can correlate functional connectivity and behavioral measures from the experiment, such as the post-error slowing and post error accuracy. In particular, our results show that frontal-central synchrony from HTS is associated with adaptive behavioral adjustments after errors, and suggest that multivariate synchrony from HTS provides the means for quantifying the functional integration of regions engaged in cognitive control.

1.3.2 Graph to Signal Transform for Weighted Graphs

In addition to bivariate and multivariate measures of functional connectivity, graph theoretic methods have been widely used in the study and characterization of functional connectivity networks. However, these methods have some drawbacks such as their dependence on node degree and shortest path distance. In Chapter 4, we introduce a method for graph to signal transform which is

applicable to both binary and weighted networks. In recent years, the relationship between graphs and signals has been exploited, providing great contributions to the analysis of complex networks. In particular, graph to signal transformations provide the means for obtaining signals that contain the structural information of the networks [43], [44], [45]. Such methods can overcome some of the drawbacks that traditional graph measures face such as the dependence on node degree and shortest path distance, and furthermore, facilitate the implementation of signal processing measures over networks by applying them directly to the signals from the networks. Applications of the graph to signal transform include the assessment of temporal networks [46] and graph filtering [47]. However, all of the previously proposed methods are only applicable to binary networks and it is not possible to apply them directly to the assessment of functional connectivity networks, which are generally weighted.

In order to apply the graph to signal transformation to functional connectivity networks, we propose to employ the resistance distance matrix as an alternative distance in the classical multi-dimensional algorithm (CMDS) [44]. We show how the proposed method serves to characterize both binary and weighted networks and its characterization analytically. Based on the signals obtained from the transformation we propose a series of approaches to study the networks. First, we propose a method for characterizing small-world networks based on the spectral centroid of the signals. As illustrated, the proposed method is more accurate in the estimation of the probability of attachment and the average degree when compared to the traditional small-world measure. In addition, we introduce graph information theoretic measures that account for the information content of the networks. Quantifying the network's entropy and divergence between networks is important as these provide insights into the network's information content [48]. In this work, we propose a graph entropy measure and a graph divergence based on the magnitude spectra of the signals from the graphs. This method is novel in the sense that the network's information is quantified from its

signals and does not depend on any graph theoretic measure nor rely on any arbitrary parameters as current graph information measures do. Finally, we propose to use the spectra of the signals from the networks in the detection of events in temporal networks. This method employs a tensor constructed from the magnitude spectra over time and uses the temporal mode of the tensor for detecting the events.

1.3.3 Dynamic Graph Fourier Transform

Following the study of the relationships between networks and signals for quantifying their structural properties, it is possible to learn from the networks by combining both the networks and the signals recorded at each node. This provides an alternative way for studying how the human brain functionally integrates the activity from various regions during cognitive processes. The recent field of signal processing over graphs [49] or graph signal processing aims to analyze signals defined over irregular domains, such as signals indexed by the nodes of a graph. This is opposed to traditional signal processing, which aims to analyze signals defined over regular domains, like time and frequency. In recent years, there has been a significant advancement in the definition of concepts widely used in traditional signal processing specifically designed for signals over graphs, such as the graph Fourier transform (GFT) and wavelet transforms, and other concepts such as sampling.

In Chapter 5, we are interested in extending the study of functional connectivity networks by considering techniques from graph signal processing. In particular, we are interested in computing the GFT of the EEG signals defined over the electrodes. However, all of the current transforms in graph signal processing consider static networks, which is not the case for functional connectivity networks which change over time depending on the underlying cognitive processes. In order to compute the graph Fourier transform on dynamic networks, a common subspace needs to be found

across time. Recently, a dynamic graph Fourier transform [50] which finds a common subspace based on Grassmann manifolds was proposed. However, the accuracy of the common subspace is compromised as the time span increases. Alternatively, in [51] the authors consider the networks averaged over time, which compromises the time-varying structure of the temporal networks.

In this chapter, we propose a dynamic Graph Fourier transform (dGFT) whose common subspace is found through the tensor decomposition of the graph Laplacian over time. By computing a dynamic graph Fourier transform over functional connectivity networks from the cognitive control experiment, it is possible to quantify the graph spectral activity from error and correct trials which suggests a higher structural organization during errors within the ERN interval when compared to correct trials.

Chapter 2

Background

In this chapter we review the background on bivariate time-frequency phase synchrony, graph theory, and describe the EEG cognitive control experiment.

2.1 Bivariate Time-Frequency Phase Synchrony

Functional connectivity networks are constructed from the bivariate time-frequency PLV, based on the Reduced Interference Rihaczek time-frequency distribution (RID-Rihaczek) as proposed in [30]. For a signal x_i , define $C_i(t, \omega)$ to be its complex RID-Rihaczek time-frequency distribution, given by

$$C_i(t, \omega) = \int \int \underbrace{\exp\left(-\frac{(\theta\tau)^2}{\sigma}\right)}_{\text{Choi-Williams kernel}} \underbrace{\exp(j\frac{\theta\tau}{2})}_{\text{Rihaczek kernel}} A_i(\theta, \tau) e^{-j(\theta t + \tau \omega)} d\tau d\theta, \quad (2.1)$$

where $A_i(\theta, \tau)$ is the ambiguity function of x_i :

$$A_i(\theta, \tau) = \int x_i(u + \frac{\tau}{2}) x_i^*(u - \frac{\tau}{2}) e^{j\theta u} du. \quad (2.2)$$

The time-varying phase of the signal x_i is computed as

$$\Phi_i(t, \omega) = \arg \left[\frac{C_i(t, \omega)}{|C_i(t, \omega)|} \right]. \quad (2.3)$$

The phase difference between two signals x_1 and x_2 can be computed as

$$\Phi_{1,2}(t, \omega) = \arg \left[\frac{C_1(t, \omega)}{|C_1(t, \omega)|} \frac{C_2^*(t, \omega)}{|C_2(t, \omega)|} \right]. \quad (2.4)$$

The PLV between two signals x_1 and x_2 as a function of time and frequency [52] is defined by

$$\begin{aligned} PLV_{1,2}(t, \omega) &= \frac{1}{N} \left| \sum_{k=1}^N \exp \left(j \Phi_{1,2}^k(t, \omega) \right) \right| \\ &= \sqrt{\langle \cos \Phi_{1,2}^k(t, \omega) \rangle^2 + \langle \sin \Phi_{1,2}^k(t, \omega) \rangle^2}, \end{aligned} \quad (2.5)$$

where N corresponds to the total number of trials or realizations of the signal, $\Phi_{1,2}^k(t, \omega)$ is the phase difference between x_1 and x_2 as defined by (2.4) for the k^{th} trial and $\langle \cdot \rangle$ denotes averaging over trials. For each trial k , the phase difference $\Phi_{1,2}^k(t, \omega)$ defines a vector on the unit circle. Thus, PLV evaluates the circular variance of the unit vectors across trials. PLV approaches 1 if the phase differences over trials exhibit small variation and approaches 0 if there is no synchrony over trials.

2.2 Graph Theory

An undirected graph $G = (V, E)$ is defined by a set of N nodes, $v_i \in V$, and a set of M edges, $e_{ij}, i, j \in \{1, \dots, N\}$. The relationships between the nodes of the graph is represented by the adja-

cency matrix $\mathbf{A} = [A_{ij}]$, for binary graphs, and $\mathbf{W} = [W_{ij}]$ for weighted graphs. In binary graphs, $A_{ij} = 1$ when nodes i and j are connected and $A_{ij} = 0$ when the nodes are not connected. For weighted graphs, W_{ij} represents the weight of the edge between nodes i and j and equals zero when $i = j$. The degree matrix Δ is defined as the diagonal matrix with entries $\Delta_{ii} = \sum_{j=1, j \neq i}^N A_{ij}$, where Δ_{ii} is the degree of node v_i . Similarly, the degree matrix Δ^w for weighted networks has diagonal entries $\Delta_{ii}^w = \sum_{j=1, j \neq i}^N W_{ij}$.

For binary graphs, the combinatorial Laplacian \mathbf{L} is defined as $\mathbf{L} = \Delta - \mathbf{A}$. The entries of \mathbf{L} are given by

$$L_{ij} = \begin{cases} \Delta_{ii}, & i = j, \\ -1, & (i, j) \in E, \\ 0, & \text{otherwise,} \end{cases} \quad (2.6)$$

where Δ_{ii} is the degree of node v_i . Similarly, for weighted graphs the Laplacian is defined as $\mathbf{L}^w = \Delta^w - \mathbf{W}$.

Another important matrix in graph theory is the incidence matrix \mathbf{C} , which is a $N \times M$ matrix, where N is the total number of nodes and M is the total number of edges in the graph. For undirected graphs, the entries C_{ij} are equal to 1 in the case that vertex v_i and edge e_j are incident and equal to zero otherwise.

The network small-worldness for binary networks is given by $\sigma = \frac{C/C_{rand}}{L/L_{rand}}$ and for weighted networks as $\sigma^w = \frac{C^w/C_{rand}^w}{L^w/L_{rand}^w}$ [10], [53]. C and C_{rand} are the clustering coefficients of the network and a random network, respectively. Similarly, L and L_{rand} are the characteristic path lengths for the analyzed network and a random network, respectively.

2.3 Cognitive control experiment

In this section, we describe the EEG dataset from the cognitive control experiment assessed in this thesis.

2.3.1 Participants

Data from nineteen subjects, all native English speakers and undergraduate students from Michigan State University, were extracted from a study of the relationship between the ERN and individual differences [54]. Subjects participated for course credit.

2.3.2 Experiment

The experiment consisted of a letter version of the Eriksen flanker task [55] which involved correctly identifying the target letter, located at the center of a five-letter string. The target was either congruent (e.g., M M M M M) or incongruent (e.g., N N M N N) with the flanker letters. Subjects pressed a determined mouse button to identify the center letter. The total time for each trial was 135 ms. Flanker letters were presented 35 ms before the target-letter onset and then the five letters remained on the screen for 100 ms. During the inter-trial intervals, ranging from 1,200–1,700 ms, a fixation cross was presented. The experiment consisted of six blocks of 80 trials and the letters comprising the strings differed between blocks. Also, the mouse button assigned to each target letter was reversed at the middle of each block.

2.3.3 EEG Data Acquisition and Pre-processing

Electroencephalographic activity was recorded by the ActiveTwo system (BioSemi, Amsterdam, The Netherlands). 64 Ag-AgCl electrodes were embedded in a stretch Lycra cap according to the

10/20 system. Also, two electrodes were located on the left and right mastoids. Electrooculogram activity generated by eye movements and blinks was recorded at electrode FP1 and three electrodes located underneath the left pupil and on the left and right outer canthi. All signals were sampled at 512 Hz by the BioSemi's ActiView software. Offline analyses were completed on BrainVision Analyzer 2 (BrainProducts, Gilching, Germany). Electrode recordings were re-referenced to the mean of the mastoids and then band-pass filtered between 0.1 and 30 Hz, with 12 dB/octave roll off. Eye movement artifacts were corrected by the regression method provided by [56]. Epochs of response-locked signals were taken 200 ms preceding the response onset and the subsequent 800 ms. Trials with a voltage difference greater than 200V within it, a voltage step greater than 50V between adjacent sampling points or voltage difference less than 0.5 mV within it were rejected. EEG signals were processed using the Current Source Density Toolbox (CSD) for volume conduction correction. The number of error trials ranged from 20 to 61 (36.78 ± 13.72 , meanst.dev.) and the same number of correct responses, per subject, were chosen randomly for the current analyses.

Chapter 3

Hypertorus Multivariate Phase Synchrony

3.1 Introduction

Coordinated time-varying interactions are fundamental in dynamical systems, ranging from a few coupled elements to complex networks. Examples of systems of coupled oscillators occur widely in nature and engineering such as circadian rhythms [57], neuroscience [58], flashing fireflies [59], coupled Josephson junctions [60], the Millenium Bridge [61], and others [62], [63], [64], [65]. In the stochastic sense, synchronization has been defined as an adjustment of rhythms of oscillating objects due to their weak interaction [66] and this adjustment can be described in terms of phase locking and frequency entrainment. Phase locking or phase synchrony between two oscillators occurs when the generalized phase difference, $\Phi_{i,j}(t, \omega) = |\Phi_i(t, \omega) - \Phi_j(t, \omega)| \bmod 2\pi < \text{constant}$, at time t and frequency ω [67], [68]. Two steps are needed for quantifying phase synchrony. First, instantaneous phase of each signal is estimated at a particular frequency of interest through methods such as the Hilbert transform, complex wavelet transform [69], empirical mode decomposition [42], [70], [71], [72], [73] or the recently proposed RID-Rihaczek complex time-frequency distribution [30], [74]. In the second step, the amount of synchrony is quantified through either the entropy of the distribution of the phase differences or mean phase coherence, also known as PLV2.1, which computes the circular variance of the relative phase [28], [75]. Although bivariate

PLV has been widely used, it has various disadvantages for the study of large and complex networks. First, PLV does not provide information about the common integrating structure among the ensemble of oscillators. Second, for large data sets multiple computations of pairwise PLV increase computational costs.

Recently, phase synchronization of a group of oscillators, which is referred to as global or multivariate phase synchronization, has been of interest for understanding the group dynamics and characteristic behavior of complex networks [72], [76], [77], [78], [79]. Contrary to the bivariate phase synchrony, multivariate synchrony captures the global synchronization patterns quantifying the degree of interactions within a group of oscillators. In addition, multivariate synchrony methods provide a single number, rather than a matrix of pairwise synchrony values. One of the earliest approaches to multivariate synchrony analysis was global field synchronization (GFS) proposed by Koenig et al. [77]. GFS first transforms the time series data to the frequency domain and then quantifies the scatter of the multivariate data through the eigenvalues of the covariance matrix of the sine and cosine coefficients of the Fourier transform. This measure inherently assumes the stationarity of the data and cannot capture time-varying aspects of synchrony. Moreover, this method quantifies synchrony as the instance when the phases of the two signals are exactly the same and does not take into account the case of constant phase difference. Knyazeva et al. [80] proposed another simple measure, the multivariate phase synchrony (MPS), defined as the mean phase synchrony averaged across the observation samples. Rudrauf et al. [81], on the other hand, proposed an alternative approach to quantifying phase synchrony through frequency locking by exploiting the relationship between phase and frequency and identifying continuous periods of identical instantaneous frequency. Similarly, in [72] the idea of cointegration is used to define multivariate phase synchrony. However, this method can only identify phase synchrony in a nonstatistical sense and is not reliable in the case of noisy signals.

More recently, methods inspired by random matrix theory (RMT) and spectral graph theory were proposed. These methods first compute the bivariate synchrony and then perform cluster analysis through eigendecomposition of the bivariate synchrony matrix as proposed by Allefeld et al. [82]. Initial work in this area focused on perceiving the oscillators as constituting a single cluster to which they participate in different degrees [83]. The existence of a single synchronization cluster is not a reasonable assumption since most complex networks usually consist of multiple clusters. In order to address this limitation, approaches based on the eigenvalue decomposition of the pairwise bivariate synchronization matrix have been proposed [84], [85]. However, it has recently been shown in cases where there are clusters of similar strength that are slightly synchronized with each other, the assumed one-to-one correspondence between eigenvectors and clusters is not realistic [86].

In order to capture the connectivity structure with a single number, Saito et al. [87] quantified global synchrony through the entropy of the eigenspectrum of the covariance or bivariate connectivity matrix. This measure was then generalized by Stam et al. [35] and others as the S-measure [36], [76], [88]. This measure uses the principle of time-delay embedding and indicates how strongly channel k at a given time is synchronized to all other channels. Similar to other methods in nonlinear dynamics, it requires the selection of different parameters, such as a threshold and the time-lag, and is computationally expensive.

Recently, HPS was introduced as an alternative method to directly measure the multivariate phase synchronization among a group of oscillators [37], [38]. HPS generalizes bivariate synchrony, where the phase difference between two time series is mapped onto a unit circle, by mapping the $N - 1$ phase differences between consecutive oscillators onto an N -dimensional space parameterized by hyperspherical coordinates [37]. HPS is advantageous over the S-estimator thanks to its reduced computational complexity and robustness to noise [38]. However, as we show in Sec-

tion 3.3, we found that this estimator was highly dependent on the ordering of the phase differences parameterizing the hypersphere.

In this chapter, we propose a novel measure to estimate the multivariate phase synchrony in a hyperdimensional coordinate system and address the shortcomings of HPS. Two complementary approaches are developed to quantify the circular variance of phase differences among multiple oscillators in a high dimensional space. In the first approach, we extend the hyperspherical coordinate system used in HPS to include redundancies, i.e. x and y coordinates of circles with varying radii, such that the ordering of the phases is not important. In the second approach, we propose a new mapping of the phase differences to a high-dimensional flat torus and compute the magnitude of the mean phase vector in this new geometry resulting in the hypertorus phase synchrony (HTS). We then show the equivalence of these two metrics, provide analytical bounds on the bias and variance of HTS and show bias correction for HTS squared. We compared the performance of HTS and the S-estimator on simulated networks of chaotic oscillators for sensitivity to coupling strength and network structure.

3.2 Background

3.2.1 S-estimator

The S-estimator at time t and frequency ω is computed as

$$S(t, \omega) = 1 + \frac{\left(\sum_{m=1}^M \lambda_m \log(\lambda_m) \right)}{\log(M)}, \quad (3.1)$$

where λ_m , $m = 1, \dots, M$ are the eigenvalues of the bivariate synchronization matrix $\{PLV_{i,j}(t, \omega)\}$, $i, j = 1, \dots, M$, and M is the total number of oscillators in the network [35, 36]. S-estimator is

equivalent to 1 minus the entropy of the normalized eigenvalues of the PLV matrix. This measure equals to 1 when all the oscillators are pairwise highly synchronized. In that case, all the entries in the PLV matrix will be equal to one and thus only one eigenvalue will be equal to one. On the other hand, when all the oscillators in the network are not pairwise synchronized the PLV matrix is full rank and its eigenvalues are uniformly distributed, maximizing the entropy and resulting in zero multivariate synchrony.

3.3 Hyperspherical Phase Synchrony

In this section, we describe the problem of HPS, which was published on [89]. Bivariate phase synchrony is based on the circular variance of the two-dimensional direction vectors on a unit circle (1-sphere), obtained by mapping the phase differences $\{\Phi_{1,2}^k(t, \omega)\}_{k=1, \dots, N}$, where N is the total number of trials, between the two time-series onto a Cartesian coordinate system. If the circular variance of these direction vectors is low, the time-series are said to be locked to each other.

HPS proposed in [37] is an extension of this idea to the multivariate case. Define

$$\theta_1^k(t, \omega), \theta_2^k(t, \omega), \dots, \theta_{M-1}^k(t, \omega) \quad (3.2)$$

as the $(M - 1)$ angular coordinates at time t and frequency ω for the k^{th} trial, where $\theta_i^k(t, \omega) = \Phi_i^k(t, \omega) - \Phi_{i+1}^k(t, \omega)$ is the phase difference between the i^{th} and $(i + 1)^{th}$ time series within a group of M oscillators. These $(M - 1)$ angular coordinates are mapped onto an M -dimensional space by forming direction vectors in an M -dimensional hyperspherical coordinate system. For any natural number M , an $(M - 1)$ -sphere of radius r is defined as the set of points in (M) -dimensional

Euclidean space which are at distance r from a central point, where the radius r may be any positive real number. The set of coordinates in an M -dimensional space, $\gamma_1, \gamma_2, \dots, \gamma_M$, that define an $(M-1)$ -sphere is represented by

$$r^2 = \sum_{i=1}^M (\gamma_i - c_i)^2, \quad (3.3)$$

where $\mathbf{c} = [c_1, \dots, c_M]$ is the center point and r is the radius. In [37], $r = 1$ and the center point is the origin.

Using the $(M-1)$ angular coordinates, a direction vector $\Gamma^k(t, \omega) = [\gamma_1^k(t, \omega), \dots, \gamma_M^k(t, \omega)]$ can be formed by mapping the angular coordinates $(\theta_1, \dots, \theta_{M-1})$ on a unit $(M-1)$ -sphere as:

$$\begin{aligned} \gamma_1^k(t, \omega) &= \cos(\theta_1^k(t, \omega)), \\ \gamma_2^k(t, \omega) &= \sin(\theta_1^k(t, \omega)) \times \cos(\theta_2^k(t, \omega)), \\ \gamma_3^k(t, \omega) &= \sin(\theta_1^k(t, \omega)) \times \sin(\theta_2^k(t, \omega)) \times \cos(\theta_3^k(t, \omega)), \\ &\vdots \\ \gamma_{M-1}^k(t, \omega) &= \sin(\theta_1^k(t, \omega)) \times \dots \times \sin(\theta_{M-2}^k(t, \omega)) \times \cos(\theta_{M-1}^k(t, \omega)), \\ \text{and } \gamma_M^k(t, \omega) &= \sin(\theta_1^k(t, \omega)) \times \dots \times \sin(\theta_{M-2}^k(t, \omega)) \times \sin(\theta_{M-1}^k(t, \omega)). \end{aligned} \quad (3.4)$$

Based on this mapping HPS is defined as

$$HPS(t, \omega) = \frac{1}{N} \left\| \sum_{k=1}^N \Gamma^k(t, \omega) \right\|_2, \quad (3.5)$$

where $HPS(t, \omega)$ is the multivariate synchronization value at time t and frequency ω , $\|\cdot\|_2$ is the Euclidean norm and N is the number of trials. In the case of perfect multivariate phase synchro-

nization of the network, HPS is equal to 1 and it equals 0 when the oscillators are independent. Note that HPS is equivalent to PLV for a network consisting of two signals. In this case, $M = 2$ and from (3.4) the direction vector $\Gamma^k(t, \omega) = [\gamma_1^k(t, \omega), \gamma_2^k(t, \omega)]$, where $\gamma_1^k(t, \omega) = \cos(\theta_1^k(t, \omega))$ and $\gamma_2^k(t, \omega) = \sin(\theta_1^k(t, \omega))$. Hence, (3.5) is equivalent to (2.5).

It can be shown that the HPS defined based on the coordinate system in (3.4) is dependent on the ordering of the phase differences $\theta_i(t, \omega)$. This dependency will result in unstable HPS values and lead to incorrect interpretation of the multivariate synchrony. To illustrate this problem, we show the derivation of the HPS value for the case of three oscillators ($M = 3$). The rotating vectors in (3.4) can be written as,

$$\begin{aligned}\gamma_1^k(t, \omega) &= \cos(\theta_1^k(t, \omega)), \\ \gamma_2^k(t, \omega) &= \sin(\theta_1^k(t, \omega)) \times \cos(\theta_2^k(t, \omega)), \\ \gamma_3^k(t, \omega) &= \sin(\theta_1^k(t, \omega)) \times \sin(\theta_2^k(t, \omega)).\end{aligned}\tag{3.6}$$

For simplicity, we further assume that we have only two trials with angular coordinates (or phase differences) $\{\theta_1^1(t, \omega), \theta_2^1(t, \omega)\}$ and $\{\theta_1^2(t, \omega), \theta_2^2(t, \omega)\}$, respectively. The corresponding HPS given in (3.5) reduces to

$$\begin{aligned}HPS(t, \omega) &= \frac{1}{N} \sqrt{\left(\sum_{k=1}^N \gamma_1^k(t, \omega)\right)^2 + \left(\sum_{k=1}^N \gamma_2^k(t, \omega)\right)^2 + \left(\sum_{k=1}^N \gamma_3^k(t, \omega)\right)^2} \\ &= \frac{1}{2} \sqrt{2 + 2 \cos(\theta_1^1(t, \omega)) \cos(\theta_1^2(t, \omega)) + 2 \sin(\theta_1^1(t, \omega)) \sin(\theta_1^2(t, \omega)) \cos(\theta_2^1(t, \omega) - \theta_2^2(t, \omega))}.\end{aligned}\tag{3.7}$$

In order to show that HPS is dependent on the ordering of the phase differences, we recalculate

the HPS with reordered angular coordinates $\{\theta_2^1(t, \omega), \theta_1^1(t, \omega)\}$ and $\{\theta_2^2(t, \omega), \theta_1^2(t, \omega)\}$,

$$HPS(t, \omega) = \frac{1}{2} \sqrt{2 + 2 \cos(\theta_2^1(t, \omega)) \cos(\theta_2^2(t, \omega)) + 2 \sin(\theta_2^1(t, \omega)) \sin(\theta_2^2(t, \omega)) \cos(\theta_1^1(t, \omega) - \theta_1^2(t, \omega))}. \quad (3.8)$$

It is clear that (3.7) and (3.8) are not equivalent except in the case of perfect synchrony, i.e., $\theta_1^1(t, \omega) = \theta_2^1(t, \omega)$, $\theta_1^2(t, \omega) = \theta_2^2(t, \omega)$, $\theta_1^1(t, \omega) = \theta_1^2(t, \omega)$ and $\theta_2^1(t, \omega) = \theta_2^2(t, \omega)$. Therefore, the ordering of the phase differences $\theta_i^k(t, \omega)$ plays a major role in calculating the corresponding HPS values. Thus, a modification of this definition is required to address this problem. In addition, in order to capture global phase information, we will replace the previously defined pairwise phase differences for HPS by the phase difference between the phase of each oscillator and the phase of the resultant vector of the remaining oscillators [90], given by

$$\theta_i^k(t, \omega) = \Phi_i^k(t, \omega) - \arg \left\{ \sum_{\substack{m=1 \\ m \neq i}}^M \exp(j\Phi_m^k(t, \omega)) \right\}. \quad (3.9)$$

3.4 Proposed Solution

3.4.1 Hyperspherical Approach

In this section, we propose a solution to the phase ordering problem encountered in HPS. This approach is based on the analysis of the hyperspherical coordinate system given in (3.4). The coordinates in (3.4) are equivalent to x coordinates of a rotating circle with varying radii. For example, $\gamma_1^k(t, \omega)$ is the x coordinate of a vector on the unit circle at angular position $\theta_1^k(t, \omega)$, while $\gamma_2^k(t, \omega)$ is the x coordinate of a vector on a circle with radius $\sin(\theta_1^k(t, \omega))$ at angular position

$\theta_2^k(t, \omega)$. Similar analysis applies to the remaining $\theta_i^k(t, \omega)$ s. Thus, every $\gamma_i^k(t, \omega)$ is just the x coordinate of a vector on a circle with radius $r_i(t, \omega) = \prod_{j=1}^{i-1} \sin(\theta_j^k(t, \omega))$, for $i = 2, 3, \dots, M$ and with a phase $\theta_i^k(t, \omega)$. The equation for $r_i(t, \omega)$ shows that as i increases, $\gamma_i^k(t, \omega)$ will have less impact on the overall synchrony. This means that the choice of the first phase difference, $\theta_1^k(t, \omega)$, will have a high impact on the measured synchrony.

Eq. (3.4) may also be interpreted as follows. Every $\gamma_i^k(t, \omega)$ is the x projection of the y coordinate of the previous $\gamma_{i-1}^k(t, \omega)$ on the x -axis with a phase $\theta_i^k(t, \omega)$, i.e. define x and y coordinates of the rotating vector for each trial k as

$$\begin{aligned}
\gamma_{x_1}^k(t, \omega) &= \cos(\theta_1^k(t, \omega)), \\
\gamma_{y_1}^k(t, \omega) &= \sin(\theta_1^k(t, \omega)), \\
\gamma_{x_2}^k(t, \omega) &= \sin(\theta_1^k(t, \omega)) \times \cos(\theta_2^k(t, \omega)), \\
\gamma_{y_2}^k(t, \omega) &= \sin(\theta_1^k(t, \omega)) \times \sin(\theta_2^k(t, \omega)), \\
\gamma_{x_3}^k(t, \omega) &= \sin(\theta_1^k(t, \omega)) \times \sin(\theta_2^k(t, \omega)) \times \cos(\theta_3^k(t, \omega)), \\
\gamma_{y_3}^k(t, \omega) &= \sin(\theta_1^k(t, \omega)) \times \sin(\theta_2^k(t, \omega)) \times \sin(\theta_3^k(t, \omega)), \\
&\vdots \\
\gamma_{x_M}^k(t, \omega) &= \sin(\theta_1^k(t, \omega)) \times \dots \times \sin(\theta_{M-1}^k(t, \omega)) \times \cos(\theta_M^k(t, \omega)), \\
\gamma_{y_M}^k(t, \omega) &= \sin(\theta_1^k(t, \omega)) \times \dots \times \sin(\theta_{M-1}^k(t, \omega)) \times \sin(\theta_M^k(t, \omega)),
\end{aligned} \tag{3.10}$$

where the phases $\theta_i^k(t, \omega)$ s are defined as in (3.9) and the superscripts x and y refer to the projection coordinates.

We can also rewrite (3.10) as,

$$\begin{aligned}
\gamma_{x_1}^k(t, \omega) &= \cos(\theta_1^k(t, \omega)), \\
\gamma_{x_2}^k(t, \omega) &= \gamma_{y_1}^k(t, \omega) \times \cos(\theta_2^k(t, \omega)), \\
\gamma_{x_3}^k(t, \omega) &= \gamma_{y_2}^k(t, \omega) \times \cos(\theta_3^k(t, \omega)), \\
&\vdots \\
\gamma_{x_M}^k(t, \omega) &= \gamma_{y_{M-1}}^k(t, \omega) \times \cos(\theta_M^k(t, \omega)), \\
\gamma_{y_M}^k(t, \omega) &= \gamma_{y_{M-1}}^k(t, \omega) \times \sin(\theta_M^k(t, \omega)).
\end{aligned} \tag{3.11}$$

Eq. (3.11) reveals that the radius $r_i^k(t, \omega) = \prod_{j=1}^{i-1} \sin(\theta_j^k(t, \omega))$, for $i = 2, 3, \dots, M$ is just the y coordinate of the previous $\gamma_{y_{i-1}}^k(t, \omega)$. This recursive structure is the cause of the ordering problem.

To solve this problem, we propose to consider both the x and y coordinates for all oscillators.

By computing the l_2 norm, $d_i^k(t, \omega)$, of the direction vectors for each oscillator i using the coordinates $\gamma_{x_i}^k(t, \omega)$ and $\gamma_{y_i}^k(t, \omega)$, we end up with the following norms,

$$\begin{aligned}
d_1^k(t, \omega) &= 1, \\
d_2^k(t, \omega) &= \sin(\theta_1^k(t, \omega)), \\
d_3^k(t, \omega) &= \sin(\theta_1^k(t, \omega)) \times \sin(\theta_2^k(t, \omega)), \\
&\vdots \\
d_M^k(t, \omega) &= \sin(\theta_1^k(t, \omega)) \times \dots \times \sin(\theta_{M-1}^k(t, \omega)),
\end{aligned} \tag{3.12}$$

or simply $d_i^k(t, \omega) = r_i^k(t, \omega) = \prod_{j=1}^{i-1} \sin(\theta_j^k(t, \omega))$ for $i = 2, 3, \dots, M$. Thus, in order to get rid of the dependency on the phase ordering, we propose to normalize $\gamma_{x_i}^k(t, \omega)$ and $\gamma_{y_i}^k(t, \omega)$ by $d_i^k(t, \omega)$. This will result in unit radius for all i . Therefore, the modified multivariate phase synchrony measure is given by

$$HPS(t, \omega) = \frac{1}{N \times \sqrt{M}} \left\| \sum_{k=1}^N D^k(t, \omega) \right\|_2, \quad (3.13)$$

where $D^k(t, \omega) = \left[\frac{\gamma_{x_1}^k(t, \omega)}{d_1^k(t, \omega)}, \frac{\gamma_{y_1}^k(t, \omega)}{d_1^k(t, \omega)}, \dots, \frac{\gamma_{x_M}^k(t, \omega)}{d_M^k(t, \omega)}, \frac{\gamma_{y_M}^k(t, \omega)}{d_M^k(t, \omega)} \right]$. As the l_2 norm of each vector $D^k(t, \omega)$ in the above equation is equal to \sqrt{M} , to make the definition and range of HPS consistent with PLV (see (2.5)) we normalize HPS by \sqrt{M} .

The modified measure given in (3.13) can be rewritten as

$$HPS(t, \omega) = \frac{1}{N\sqrt{M}} \sqrt{\left(\sum_{k=1}^N \frac{\gamma_{x_1}^k(t, \omega)}{d_1^k(t, \omega)} \right)^2 + \left(\sum_{k=1}^N \frac{\gamma_{y_1}^k(t, \omega)}{d_1^k(t, \omega)} \right)^2 + \dots + \left(\sum_{k=1}^N \frac{\gamma_{x_M}^k(t, \omega)}{d_M^k(t, \omega)} \right)^2 + \left(\sum_{k=1}^N \frac{\gamma_{y_M}^k(t, \omega)}{d_M^k(t, \omega)} \right)^2} \quad (3.14)$$

By noting that $\cos(\theta_i^k(t, \omega)) = \frac{\gamma_{x_i}^k(t, \omega)}{d_i^k(t, \omega)}$ and $\sin(\theta_i^k(t, \omega)) = \frac{\gamma_{y_i}^k(t, \omega)}{d_i^k(t, \omega)}$ we can write the modified *HPS* as

$$\begin{aligned} HPS(t, \omega) &= \frac{1}{\sqrt{M}} \sqrt{PLV_1^2(t, \omega) + \dots + PLV_M^2(t, \omega)} \\ &= \sqrt{\frac{1}{M} \sum_{i=1}^M PLV_i^2(t, \omega)}, \end{aligned} \quad (3.15)$$

where PLV_i^2 quantifying the synchronization of each oscillator with respect to a common reference

angle with $\theta_i^k(t, \omega)$ as defined in (3.9) and PLV_i is given by

$$\begin{aligned} PLV_i(t, \omega) &= \frac{1}{N} \left| \sum_{k=1}^N \exp(j\theta_i^k(t, \omega)) \right| \\ &= \sqrt{\langle \cos \theta_i^k(t, \omega) \rangle^2 + \langle \sin \theta_i^k(t, \omega) \rangle^2}. \end{aligned} \quad (3.16)$$

The maximum value of $HPS(t, \omega)$ is 1, when there is complete phase synchronization among oscillators. On the other hand, $HPS(t, \omega)$ is theoretically 0 when the oscillators are independent.

3.4.2 Hypertorus Synchrony

Results found in the previous section can alternatively be derived from an alternative mapping: the Cartesian product of unit circles parameterized by phase differences as given in (3.9). In a network consisting of M oscillators, consider a phase $\theta_i^k(t, \omega)$ that parameterizes the unit circle $\mathbb{S}^1 \subset \mathbb{R}^2$ by the angular coordinates $\mathbb{S}^1 = \{(\cos(\theta_i^k), \sin(\theta_i^k)) \mid 0 \leq \theta_i^k \leq 2\pi\}$. Let another unit circle $\mathbb{S}^1 \subset \mathbb{R}^2$ be parameterized by the angular coordinates $\mathbb{S}^1 = \{(\cos(\theta_j^k), \sin(\theta_j^k)) \mid 0 \leq \theta_j^k \leq 2\pi\}$. The Cartesian product $\mathbb{S}^1 \times \mathbb{S}^1$ defines the manifold

$$\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1 = \left\{ (\cos(\theta_i^k), \cos(\theta_j^k), \sin(\theta_i^k), \sin(\theta_j^k)) \mid 0 \leq \theta_i^k, \theta_j^k \leq 2\pi \right\}, \quad (3.17)$$

which is embedded in \mathbb{R}^4 . The M -dimensional flat torus $\mathbb{T}^m \subset \mathbb{R}^{2m}$ is the manifold $\mathbb{T}^m = \mathbb{S}^1 \times \dots \times \mathbb{S}^1$ defined by $x_1^2 + y_1^2 = \dots = x_M^2 + y_M^2 = 1$. It is parameterized by $x_i = \cos(\theta_i(t, \omega))$ and $y_i = \sin(\theta_i(t, \omega))$ [91].

A Riemannian metric g on a n -dimensional manifold M defines an inner product between tangent vectors in each tangent space $T_p M$ for every point $p \in M$ [91]. A Riemannian manifold

(M, g) is a differentiable manifold equipped with a Riemannian metric [92]. Thus, for every point p in (M, g) the length of any tangent vector $X \in T_p M$ is given by $|X| = \langle X, X \rangle^{1/2}$ [91]. The Cartesian product between two Riemannian manifolds (M_1, g_1) and (M_2, g_2) is equipped with the product metric $g = g_1 \oplus g_2$, which is defined as [91]

$$g(X_1 + X_2, Y_1 + Y_2) = g(X_1 + Y_1) + g(X_2 + Y_2), \quad (3.18)$$

where $X_i, Y_i \in T_{p_i} M_i$ and $T_{(p_1, p_2)}(M_1 \times M_2) = T_{p_1} M_1 \oplus T_{p_2} M_2$.

A torus \mathbb{T}^m is locally isometric to Euclidean space, meaning that every point on \mathbb{T}^m has a neighborhood that is isometric to an open set in \mathbb{R}^m [91], which results in a manifold whose curvature is zero everywhere and its tangent spaces are identical to the manifold [93]. Hence, \mathbb{T}^m is a flat Riemannian manifold equipped with the Euclidean metric [94].

For a group of M oscillators, vector (3.19) lies in \mathbb{T}^m

$$\Gamma^k(t, \omega) = [x_1^k, y_1^k, \dots, x_M^k, y_M^k], \quad (3.19)$$

where $x_i^k = \cos(\theta_i^k(t, \omega))$, $y_i^k = \sin(\theta_i^k(t, \omega))$, and $\theta_i^k(t, \omega)$ is a phase difference as defined in (3.9) for the k^{th} trial.

$HTS(t, \omega)$ can then be defined as

$$HTS(t, \omega) = \frac{1}{N\sqrt{M}} \left\| \sum_{k=1}^N \Gamma^k(t, \omega) \right\|_2, \quad (3.20)$$

$$\begin{aligned}
HTS(t, \omega) &= \frac{1}{N\sqrt{M}} \sqrt{\left(\sum_{k=1}^N \cos \theta_1^k(t, \omega)\right)^2 + \left(\sum_{k=1}^N \sin \theta_1^k(t, \omega)\right)^2 + \cdots + \left(\sum_{k=1}^N \cos \theta_M^k(t, \omega)\right)^2 + \left(\sum_{k=1}^N \sin \theta_M^k(t, \omega)\right)^2} \\
&= \frac{1}{\sqrt{M}} \sqrt{\langle \cos \theta_1^k(t, \omega) \rangle^2 + \langle \sin \theta_1^k(t, \omega) \rangle^2 + \cdots + \langle \cos \theta_M^k(t, \omega) \rangle^2 + \langle \sin \theta_M^k(t, \omega) \rangle^2} \\
&= \frac{1}{\sqrt{M}} \sqrt{PLV_1^2(t, \omega) + \cdots + PLV_M^2(t, \omega)} \\
&= \sqrt{\frac{1}{M} \sum_{i=1}^M PLV_i^2(t, \omega)}, \tag{3.21}
\end{aligned}$$

where M is the number of oscillators and N is the total number of trials. HTS can be re-expressed as shown in (3.21), which is equivalent to (3.15). Throughout the rest of this article we will use HTS to refer to both approaches.

3.4.3 Computational Complexity

HTS involves the computation of M PLVs, with complexity $\mathcal{O}(M \log n)$ per time-frequency point [95], where n is the number of points used in the fast Fourier transforms (ffts) in the computation of the time-frequency distribution (usually equal or greater than the length of the signal). The computation of one square root has complexity $\mathcal{O}(m^2)$ when computed through the Newton-Raphson Method [96], where m corresponds to the minimum of the number of bits from the two numbers being multiplied (32 or 64 bits for double precision). Thus, the total computational complexity of HTS is $\mathcal{O}(M \log n) + \mathcal{O}(m^2)$. On the other hand, the computational complexity of the S-estimator relies on the computation of $\binom{M}{2}$ PLVs for the construction of the synchronization matrix and its eigendecomposition. Computing $\binom{M}{2}$ PLVs has a complexity of $\mathcal{O}(\binom{M}{2} \log n)$, which can be approximated as $\mathcal{O}(\frac{M}{2} \log n)$ for large M . The eigendecomposition of the synchronization matrix has complexity $\mathcal{O}(M^3)$ [97]. Thus, the total computational complexity of the S-estimator is $\mathcal{O}(\frac{M}{2} \log n) + \mathcal{O}(M^3)$. Therefore, the proposed metric is computationally more efficient than the S-estimator.

3.5 Statistical assessment of HTS

In this section, we assessed the asymptotic properties of the expected value and variance of $\widehat{HTS}(t, \omega)$ in the absence of synchrony as well as for different levels of synchrony. Finally, as previously done for PLV, we found an unbiased estimator for \widehat{HTS}^2 and evaluated its variance empirically.

3.5.1 Bias

Due to its dependence on PLV, the proposed measure also exhibits a bias which is dependent on the number of trials. We will first illustrate this dependency by assuming a Von Mises distribution for phase differences. The Von Mises distribution $VM(\theta, \kappa)$ is the most common model for circular data [90]. It is defined by the reference direction, θ , and its dispersion about that direction, κ . Its probability density function is given by

$$f(\theta) = [2\pi I_0(\kappa)]^{-1} \exp[\kappa \cos(\theta - \mu)],$$

$$0 \leq \theta < 2\pi, 0 \leq \kappa < \infty, \quad (3.22)$$

where $I_0(\kappa) = (2\pi)^{-1} \int_0^{2\pi} \exp[\kappa \cos(\phi - \mu)]$ is the modified Bessel function of order zero [90].

Figure 3.1 illustrates the theoretical and experimental multivariate synchrony, HTS and \widehat{HTS} , respectively, for different levels of synchronization in a network consisting of $M = 4$ oscillators. Here we are assuming that the phase differences in (3.15) and (3.21) are equally distributed according to $VM(0, \kappa)$ for simplicity and various levels of synchrony are obtained by varying the concentration parameter κ . As observed, the bias of HTS depends on the underlying distribution

of the angles θ_i , bias being the most prominent in the absence of synchrony, or when θ_i is uniformly distributed. In addition, the bias is dependent on the sample size and results based on small sample sizes should be interpreted carefully.

In this thesis, we further assessed the bounds on the bias and variance of \widehat{HTS} . A lower bound on bias can be found from the inequality for arithmetic and quadratic means [98] as

$$\frac{PLV_1(t, \omega) + \dots + PLV_M(t, \omega)}{M} \leq \sqrt{\frac{PLV_1^2(t, \omega) + \dots + PLV_M^2(t, \omega)}{M}}, \quad (3.23)$$

where the absolute value in the original inequality is no longer required since $\widehat{PLV}(t, \omega) \in [0, 1]$.

An upper bound can be found as

$$\sqrt{\frac{PLV_1^2(t, \omega) + \dots + PLV_M^2(t, \omega)}{M}} \leq \frac{PLV_1(t, \omega) + \dots + PLV_M(t, \omega)}{\sqrt{M}}. \quad (3.24)$$

Thus, the lower and upper bounds on the expected value of $\widehat{HTS}(t, \omega)$ can be found as

$$\begin{aligned} E[\widehat{HTS}(t, \omega)] &= E\left[\sqrt{\frac{1}{M} \sum_{i=1}^M \widehat{PLV}_i^2(t, \omega)}\right] \\ &\geq E\left[\frac{1}{M} \sum_{i=1}^M \widehat{PLV}_i(t, \omega)\right] \\ &= \frac{1}{M} \sum_{i=1}^M E[\widehat{PLV}_i(t, \omega)], \end{aligned} \quad (3.25)$$

and

$$E[\widehat{HTS}(t, \omega)] \leq E\left[\frac{1}{\sqrt{M}} \sum_{i=1}^M \widehat{PLV}_i(t, \omega)\right], \quad (3.26)$$

respectively.

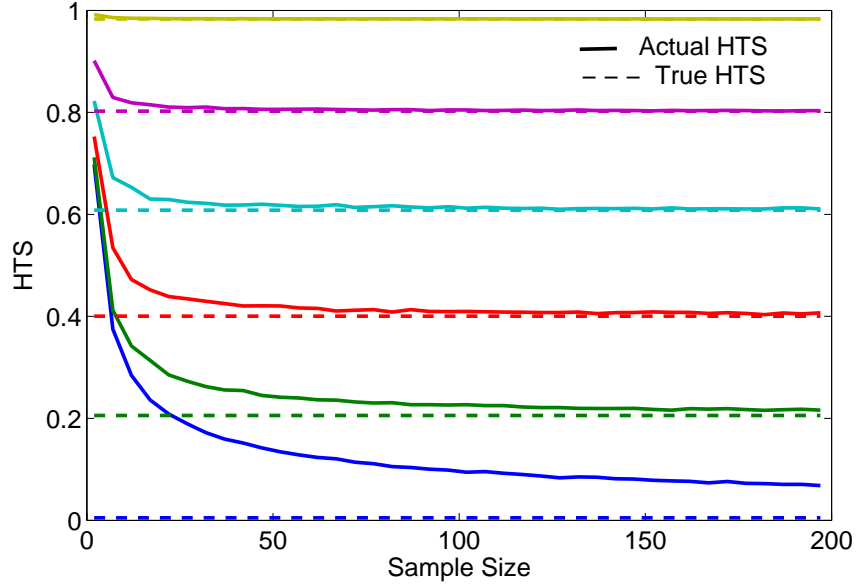


Figure 3.1: \widehat{HTS} (solid lines) and true HTS (dashed lines) for synchrony values of 0, 0.20, 0.40, 0.60, 0.80 and 0.99. $M = 4$ oscillators were simulated with phases θ_i distributed as $VM(0, \kappa)$.

An approximate value for $E \left[\widehat{PLV}(t, \omega) \right]$ in the absence of synchrony has been previously found to be $E \left[\widehat{PLV}(t, \omega) \right] \approx \frac{1}{\sqrt{N}}$ [30]. Hence, $E \left[\widehat{HTS}(t, \omega) \right]$ can be bounded as

$$\frac{1}{\sqrt{N}} \leq E \left[\widehat{HTS}(t, \omega) \right] \leq \sqrt{\frac{M}{N}}. \quad (3.27)$$

Eq. (3.27) shows that in a network in which all oscillators are independent, the minimum possible value that \widehat{HTS} can attain is inversely proportional to the square root of the total number of trials or observations, as previously found for \widehat{PLV} . On the other hand, its upper bound is directly proportional to the number of oscillators.

Asymptotic results for $E \left[\widehat{PLV}(t, \omega) \right]$, or the mean resultant length as known in the circular statistics community, have been found for the Von Mises distribution with mean direction $\theta = 0$

and concentration parameter $\kappa > 0$ [99] as

$$E \left[\widehat{PLV}(t, \omega) \right] = A(\kappa) + \frac{1}{2N\kappa}, \quad (3.28)$$

where $A(\kappa) = \frac{I_1(\kappa)}{I_0(\kappa)}$ and $I_i(\kappa)$ is the modified Bessel function of the first kind of the i^{th} order.

Considering all M oscillator's θ_i s to be identically distributed and substituting (3.28) into (3.25)

and (3.26) yields

$$\left(A(\kappa) + \frac{1}{2N\kappa} \right) \leq E \left[\widehat{HTS}(t, \omega) \right] \leq \sqrt{M} \left(A(\kappa) + \frac{1}{2N\kappa} \right). \quad (3.29)$$

3.5.2 Variance

In order to find an upper bound on the variance of $\widehat{HTS}(t, \omega)$ we will define $\widehat{HTS}^2(t, \omega)$ as

$$\widehat{HTS}^2(t, \omega) = \frac{1}{M} \sum_{i=1}^M \widehat{PLV}_i^2(t, \omega). \quad (3.30)$$

Taking expectation on both sides yields

$$\begin{aligned} E \left[\widehat{HTS}^2(t, \omega) \right] &= E \left[\frac{1}{M} \sum_{i=1}^M \widehat{PLV}_i^2(t, \omega) \right] \\ &= \frac{1}{M} \sum_{i=1}^M E \left[\widehat{PLV}_i^2(t, \omega) \right], \end{aligned} \quad (3.31)$$

where the linearity property of expectation has been employed. The expected value of $\widehat{PLV}^2(t, \omega)$

is a well known expression [34], [99], given by

$$E \left[\widehat{PLV}_i^2(t, \omega) \right] = \frac{1}{N} + \left(1 - \frac{1}{N} \right) PLV_i^2(t, \omega). \quad (3.32)$$

Thus, substituting (3.32) in (3.31) yields

$$E \left[\widehat{HTS}^2(t, \omega) \right] = \frac{1}{N} + \left(1 - \frac{1}{N} \right) HTS^2(t, \omega). \quad (3.33)$$

An upper bound on the variance of $\widehat{HTS}(t, \omega)$ in the absence of synchrony can be found as

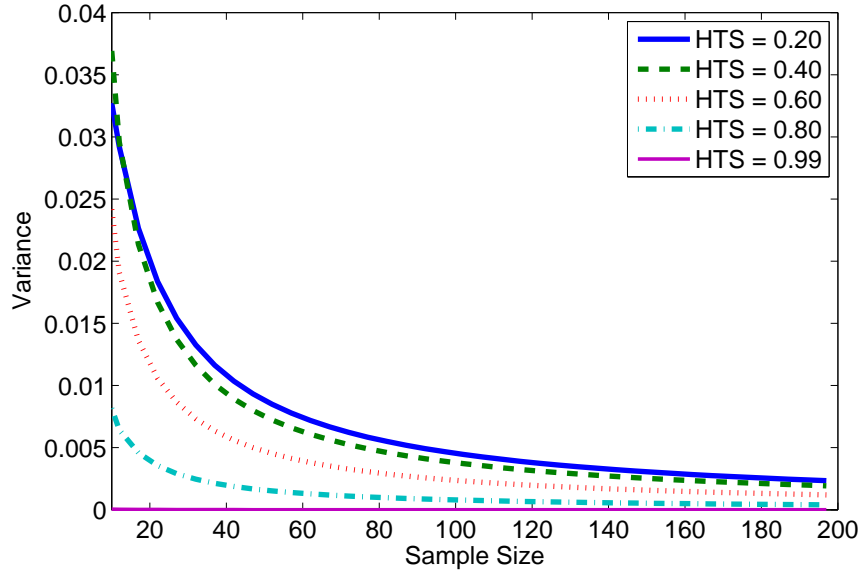
$$\begin{aligned} Var \left(\widehat{HTS}(t, \omega) \right) &= E \left[\widehat{HTS}^2(t, \omega) \right] - E \left[\widehat{HTS}(t, \omega) \right]^2 \\ &\leq \frac{1}{N} + \left(1 - \frac{1}{N} \right) HTS^2(t, \omega) - \left(\frac{1}{\sqrt{N}} \right)^2 \\ &= 1 - \frac{1}{N} HTS^2(t, \omega). \end{aligned} \quad (3.34)$$

Thus, in the absence of synchrony, the maximum possible value that the variance of \widehat{HTS} can attain is 1.

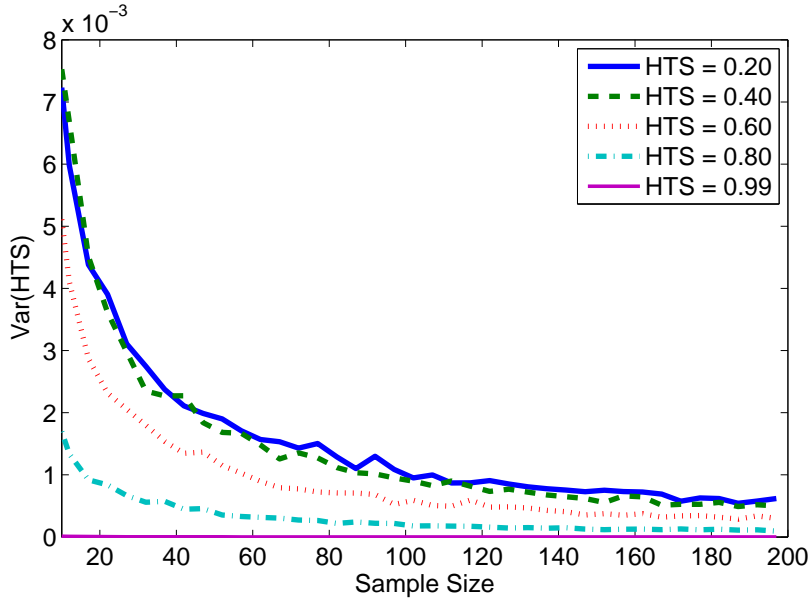
In the case of $VM(0, \theta)$ the upper bound for the variance is

$$Var \left(\widehat{HTS}(t, \omega) \right) \leq \left(\frac{1}{N} + \left(1 - \frac{1}{N} \right) HTS^2(t, \omega) \right) - \left(A^2(\kappa) + 2A(\kappa) \frac{1}{2N\kappa} + \frac{1}{4N^2\kappa^2} \right). \quad (3.35)$$

where phase differences θ_i are drawn from $VM(0, \kappa)$.



(a) Upper bounds for variance of $\widehat{HTS}(t, \omega)$.



(b) Empirical variance of $\widehat{HTS}(t, \omega)$.

Figure 3.2: (a) Theoretical upper bounds for the variance of HTS; and (b) empirical variance of HTS as a function of sample size in a network of $M = 4$ oscillators for different synchronization levels in the Von Mises distribution in (3.35).

Figure 3.2 (a) and (b) show the upper bounds on the variance of $\widehat{HTS}(t, \omega)$ and its empirical variance, respectively, for various levels of synchrony in a network consisting of $M = 4$ oscillators.

From Figure 3.2 (a), the variance of $\widehat{HTS}(t, \omega)$ decreases as the number of trials or observations increase as well as when the global synchronization increases. Figure 3.2 (b) shows the variance of $\widehat{HTS}(t, \omega)$ obtained empirically for various levels of synchrony. It is observed that the empirical variance follows similar trends as those obtained by the upper bound, without attaining it.

3.5.3 Correction of Bias in \widehat{HTS}^2

As in the case of PLV [34, 100], it is straightforward to find an unbiased estimator of $\widehat{HTS}^2(t, \omega)$ rather than for $\widehat{HTS}(t, \omega)$. Eq. (3.33) suggests that the bias in \widehat{HTS} arises from the bias in \widehat{PLV} . An unbiased estimator for $\widehat{HTS}^2(t, \omega)$ can be found as

$$\widehat{HTS}_{UB}^2(t, \omega) = \frac{1}{N-1} \left(\widehat{HTS}^2(t, \omega) \times N - 1 \right). \quad (3.36)$$

This result is obtained similarly by substituting $\widehat{PLV}^2(t, \omega)$ in (3.31) by its unbiased estimator previously found in [34], [100]

$$\begin{aligned} \widehat{PLV}_{i(UB)}^2(t, \omega) &= \frac{2}{N(N-1)} \sum_{n=1}^{N-1} \sum_{m=n+1}^N \cos(\phi_i^m(t, \omega) - \phi_i^n(t, \omega)) \\ &= \frac{1}{N-1} \left(\widehat{PLV}_i^2(t, \omega) \times N - 1 \right). \end{aligned} \quad (3.37)$$

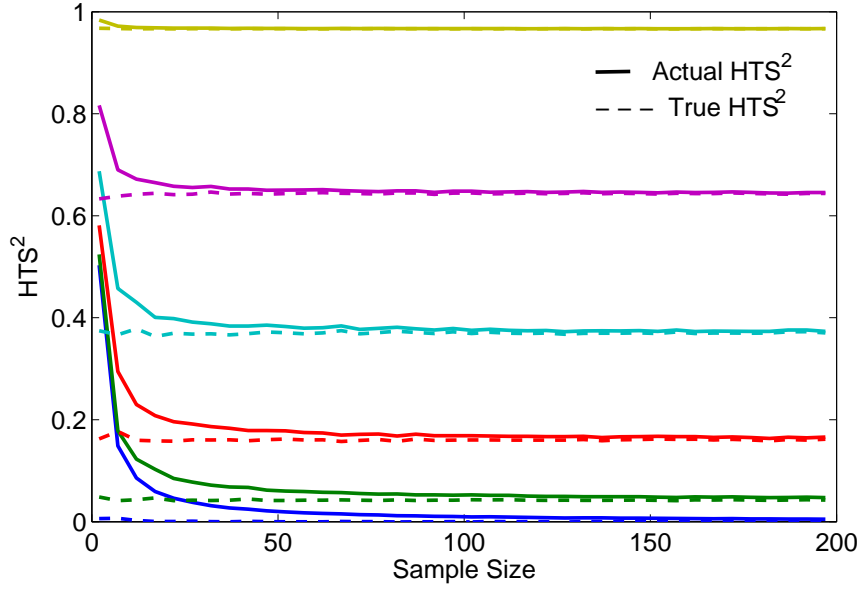


Figure 3.3: \widehat{HTS}^2 (solid lines) and true HTS^2 (dashed lines) for true HTS values of 0, 0.20, 0.40, 0.60, 0.80 and 0.99.

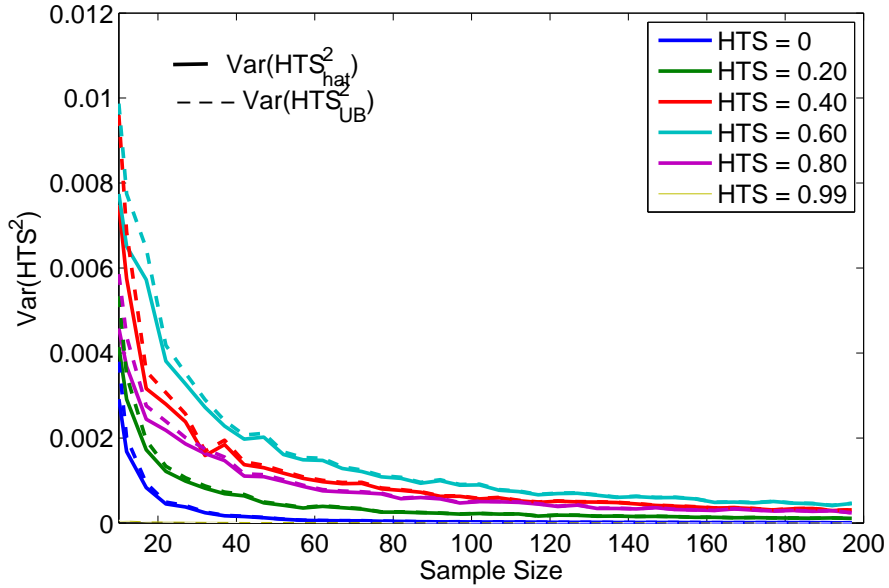


Figure 3.4: Variance of \widehat{HTS}^2 (solid lines) and true HTS^2 (dashed lines) for true HTS values of 0, 0.20, 0.40, 0.60, 0.80 and 0.99.

Figs. 3.3 and 3.4 illustrate the expected value and variance of $\widehat{HTS}^2(t, \omega)$ and $\widehat{HTS}_{UB}^2(t, \omega)$ for various synchrony values. As previously reported for \widehat{PLV}^2 , the variance of the unbiased estimator

is slightly higher than that of the biased estimator for small sample sizes.

3.6 Results

In this section, results of multivariate phase synchronization on simulated and EEG data are presented. First, the robustness to noise of proposed measure is evaluated on a network of highly synchronized sinusoidal oscillators. Next, we assess the effect of the number of oscillators on the synchrony measures. In addition, we evaluate the multivariate synchrony of a network of Kuramoto oscillators with varying coupling strengths. Next, the sensitivity to global coupling is evaluated for various Rössler oscillators, followed by the assessment of the topographical sensitivity of the proposed measure. Finally, we present a detailed analysis of global connectivity in the cognitive control experiment.

3.6.1 Assessment of robustness to noise of multivariate phase synchrony measures

In the first simulation, the performance of HTS was evaluated for various signal-to-noise ratio (SNR) levels in a network of synchronized cosine oscillators and compared to S-estimator. A network of 8 sinusoidal oscillators with constant phase differences is defined as

$$x_i(t) = \cos\left(80\pi t + \pi \frac{(i-1)}{8}\right) + \eta_i(t), \quad i = 1, \dots, 8, \quad t = 1, 2, \dots, 512, \quad (3.38)$$

where $\eta_i(t)$ is independent white Gaussian noise, with SNR between -20 and 30 dB in steps of 2 dB, for a total of N=200 trials and the signal length equal to $T = 512$ samples. S-estimator and HTS were computed by first estimating the instantaneous phase at the frequency of interest, 40 Hz,

through Hilbert transform and then evaluating the synchrony through (3.1) and (3.20). S-estimator and HTS were then averaged over time to obtain $H\bar{T}S = \frac{1}{T} \sum_{t=1}^T HTS(t)$ and $\bar{S} = \frac{1}{T} \sum_{t=1}^T S(t)$, where $T = 512$.

Figure 3.5 shows the average multivariate synchrony, for HTS and S-estimator as a function of SNR in the network of sinusoidal oscillators given by (3.38). Both, HTS and S-estimator result in multivariate synchrony equal to 1 for high SNR. However, HTS is more robust to noise as the S-estimator shows a sharper decrease for SNR values less than 10 dB. At low SNR, S-estimator approaches 0 whereas HTS remains constantly high due to bias in the PLV estimator [30], [34], [100].

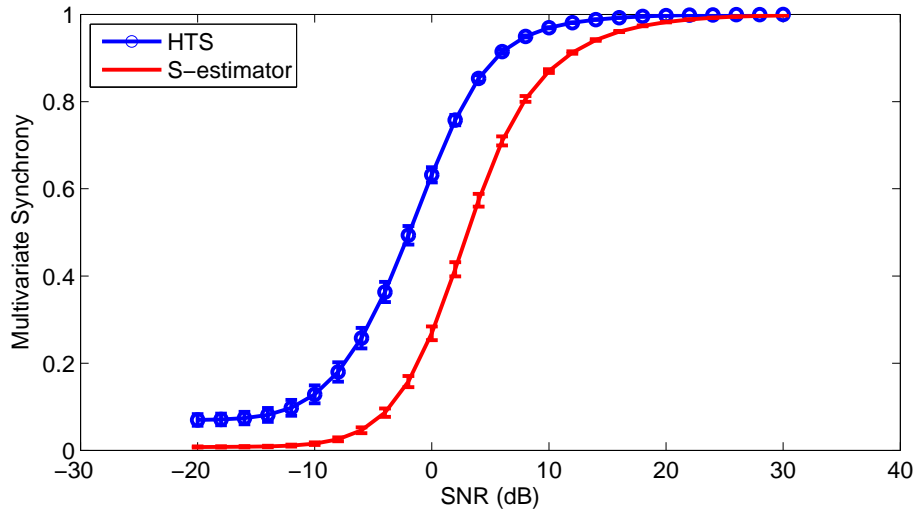


Figure 3.5: Multivariate synchrony for a network of highly synchronized sinusoidal oscillators.

3.6.2 Effect of number of oscillators on the multivariate synchrony measures

In order to show how the number of oscillators in a network affects the accuracy of the estimators, S-estimator was computed for four simulated bivariate synchronization matrices whose off-diagonal entries were uniformly distributed between $[0.05, 0.15]$, $[0.25, 0.35]$, $[0.45, 0.55]$ and

[0.65, 0.75]. The HTS was computed by letting the PLVs in 3.16 to take values from the same four intervals as for the S-estimator. These simulation were repeated 100 times. The number of oscillators in the network was incremented in powers of 2, starting at 4 up to 256. In a network where all of the oscillators have similar bivariate pairwise synchrony, we would expect the multivariate synchrony to be around that value. As shown on Figure 3.6, the multivariate synchrony obtained from the S-estimator varies depending on the number of oscillators in the network. In general, the S-estimator results in a lower synchronization value than expected, where the largest differences occur at higher synchrony. HTS, on the other hand, is not affected by the number of oscillators since it is computed through the l_2 -norm and gives synchrony values close to the true value.

In order to understand the sensitivity of the S-estimator to the number of oscillators, Figure 3.7 shows the normalized eigenvalues of a synchronization matrix whose off-diagonal entries are close to 0.4, for 4, 8, 12 and 16 oscillators. As Figure 3.7 suggests, increasing the number of oscillators results in a reduction of the entropy. Also, as the number of oscillators increases the largest normalized eigenvalue decreases. These findings show how the S-estimator is affected by the number of oscillators and results in lower multivariate synchrony than the true global synchronization in the network, potentially due to approximating the entropy through a finite number of samples. These results suggest that multivariate synchrony using the S-estimator could be underestimated unless the total number of oscillators in the network is high.

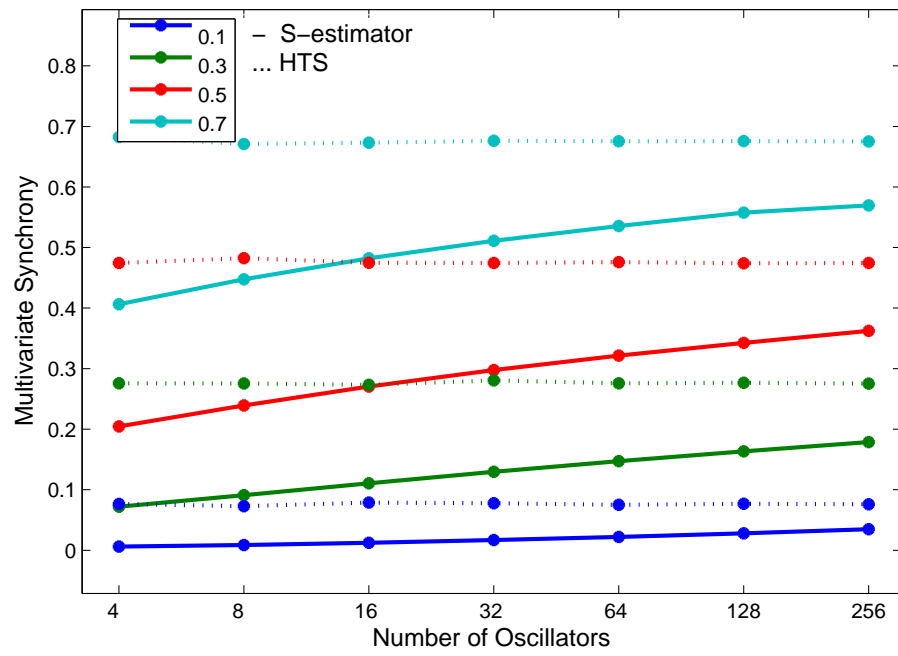


Figure 3.6: Effect of number of oscillators on S-estimator and HTS for mean synchrony values of 0.1, 0.3, 0.5, and 0.7.

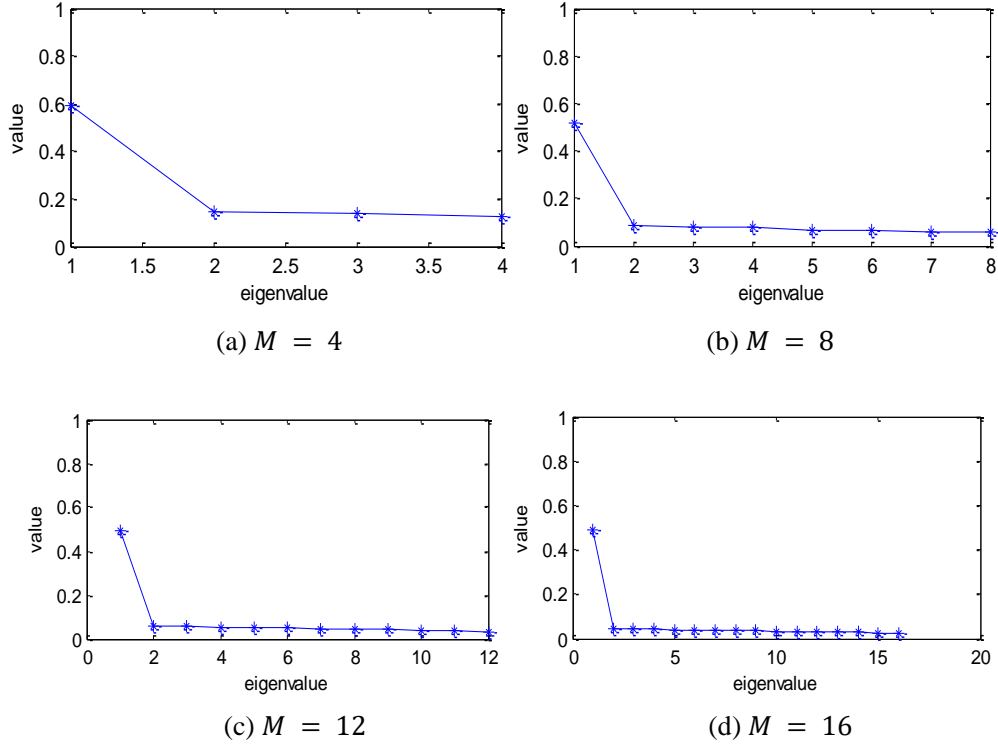


Figure 3.7: Effect of number of oscillators (M) on the eigenvalues for a true multivariate synchrony value of 0.4, a) $M = 4$; b) $M = 8$; c) $M = 12$; and d) $M = 16$.

3.6.3 Kuramoto Model

In order to evaluate the performance of the proposed measure as a function of coupling strength, we computed the multivariate synchrony in a large network of coupled oscillators as presented by Kuramoto [101]. Kuramoto model describes a system consisting of multiple oscillators with different natural frequencies which synchronize to a common frequency after their coupling exceeds a certain threshold [102]. This model has been used to describe many physical phenomena, ranging from unicellular organisms [103] to the neurosciences [74], [104]. Phase dynamics governing the cooperative synchronization among M oscillators are given by

$$\frac{d\phi_i}{dt} = \omega_i + \frac{K}{M} \sum_{j=1}^M \sin(\phi_j - \phi_i), \quad (3.39)$$

where ϕ_i corresponds to the phase of the i th oscillator, ω_i is its natural frequency and K corresponds to the coupling strength, which is equal among all oscillators. The natural frequency of each oscillator is chosen randomly from a Lorentzian distribution given by

$$g(\omega) = \frac{\gamma}{\pi [\gamma^2 + (\omega - \omega_o)^2]}, \quad (3.40)$$

with mean ω_o and width γ .

Kuramoto found that oscillators are desynchronized until K exceeds a critical value $K_c = 2\gamma$. Exceeding K_c separates the oscillators into two groups: one that contributes to the synchronization of the system and another whose natural frequencies come from the tails of the distribution and contribute to desynchronization of the system [105]. As K increases, the group of synchronized oscillators increases until all oscillators are synchronized. A network consisting of $M = 64$ oscillators was simulated and the time-varying phases $\phi_i(t)$ were solved numerically via Runge-Kutta with a time step of $\Delta t = 0.0078$ s, which results in a sampling frequency of 128 Hz. The natural frequencies of each oscillator are drawn from a Lorentzian distribution as given by (3.40) where $\omega_o = 40$ rad/s and $\gamma = 1$. This results in a $K_c = 2\gamma = 2$. The signal length was 2048 samples, and the first 500 samples were discarded to avoid transients.

Figure 3.8 shows multivariate synchrony estimated from HTS and the S-estimator as K increased from 0 to 9 in increments of 0.5. We expect to observe low synchrony for $K < K_c$ with a sudden increase in synchrony after K_c . When $K = 0$ multivariate synchrony from both S-estimator and HTS is greater than 0, which indicates bias on the estimators when phases come from an uniform distribution. On the other hand, HTS is more sensitive to the increase of global synchronization for $K = K_c = 2$ compared to the S-estimator. The standard deviation of both estimators is maximal around K_c [104], with S-estimator showing less variance than HTS since it is a weighted

average of all bivariate PLVs, obtained from the eigendecomposition. Finally, when the system is fully synchronized HTS approaches 1 as expected.

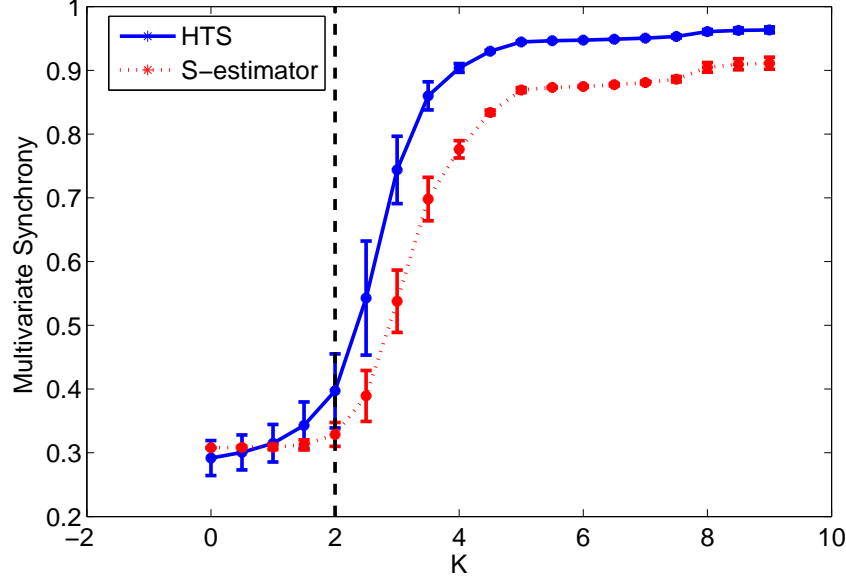


Figure 3.8: Comparison of mean and standard deviation of multivariate synchrony (HTS and S-estimator) within a Kuramoto network with $K_c = 2$.

3.6.4 Rössler oscillator

In order to test multivariate synchrony under different network configurations we used a Rössler oscillator model. Rössler oscillators describe a system of weakly coupled self-sustained stochastic oscillators [106]. We modeled a network consisting of 6 oscillators coupled through their x -dimension [107]. Eight different configurations are considered, illustrated in Fig 3.9. It is expected that networks 1 and 2 will exhibit low synchrony, and network 8 will result in multivariate synchrony close to 1. Dynamics governing the networks under study are given by

$$\begin{aligned}
\dot{\xi}_j &= \begin{pmatrix} \dot{X}_j \\ \dot{Y}_j \\ \dot{Z}_j \end{pmatrix} \\
&= \begin{pmatrix} \dot{X}_j = -\omega_j Y_j - Z_j + \left[\sum_{i \neq j} \varepsilon_{ij} (X_i - X_j) \right] + \sigma \eta_j \\ \dot{Y}_j = -\omega_j X_j - a Y_j \\ \dot{Z}_j = b + (X_j - c) Z_j \end{pmatrix},
\end{aligned} \tag{3.41}$$

where $i, j = 1, 2, \dots, 6$, $a = 0.35$, $b = 0.2$, $c = 10$, $\omega_1 = 1.05$, $\omega_2 = 1.03$, $\omega_3 = 1.01$, $\omega_4 = 0.99$, $\omega_5 = 0.97$, $\omega_6 = 0.95$, $\varepsilon_{ij} = \varepsilon_{ji} = 0.5$, $\sigma = 1.5$ and η_j is white Gaussian noise. The differential equations were solved by the Runge-Kutta method at a time step of 0.067 seconds. Simulations were repeated 200 times, for a signal length of 2000 samples and sampling frequency of 15 Hz.

Table 1 compares multivariate synchrony evaluated using HTS and S-estimator for each of the eight Rössler networks presented in Figure 3.9. The second and third columns show results for HTS and S-estimator (mean \pm st.dev.) computed according to (3.20) and (3.1), respectively. Multivariate synchrony values obtained from both measures are comparable and align with our expectations for all networks. For both methods, the multivariate synchrony results for each network is significantly different from that obtained from a null network in which none of the oscillators is connected, i.e. $\varepsilon_{ij} = 0$, (Wilcoxon rank sum test, $p < 0.01$).

The two networks differ in their behavior only for networks 5 and 6. In the case of network 5, multivariate synchrony obtained from HTS is higher than that from network 6, whereas it is the opposite for S-estimator. In network 5, four out of six oscillators are all interconnected with only

two isolated oscillators contributing to low synchrony. Since HTS relies on the root mean square of PLVs with one PLV computed between each oscillator and the mean phase, there will be only two PLVs with low synchrony. On the other hand, in network 6 although there are two sub-networks that are fully synchronized these are not interconnected and hence the global synchrony of the network should not be as high as in network 5 as indicated by HTS. This result is also observed from the unbiased squared HTS and S-estimator, as shown in the fourth and fifth columns of Table 1, respectively. Here, HTS_{UB}^2 is computed as in (3.36) and S_{UB}^2 is obtained by using unbiased PLV^2 as in (3.37). Note that network 4 also contains 6 connections and results in higher synchrony than networks 5 and 6. This is due to the indirect connections that emerge when oscillators are interconnected through a third oscillator.

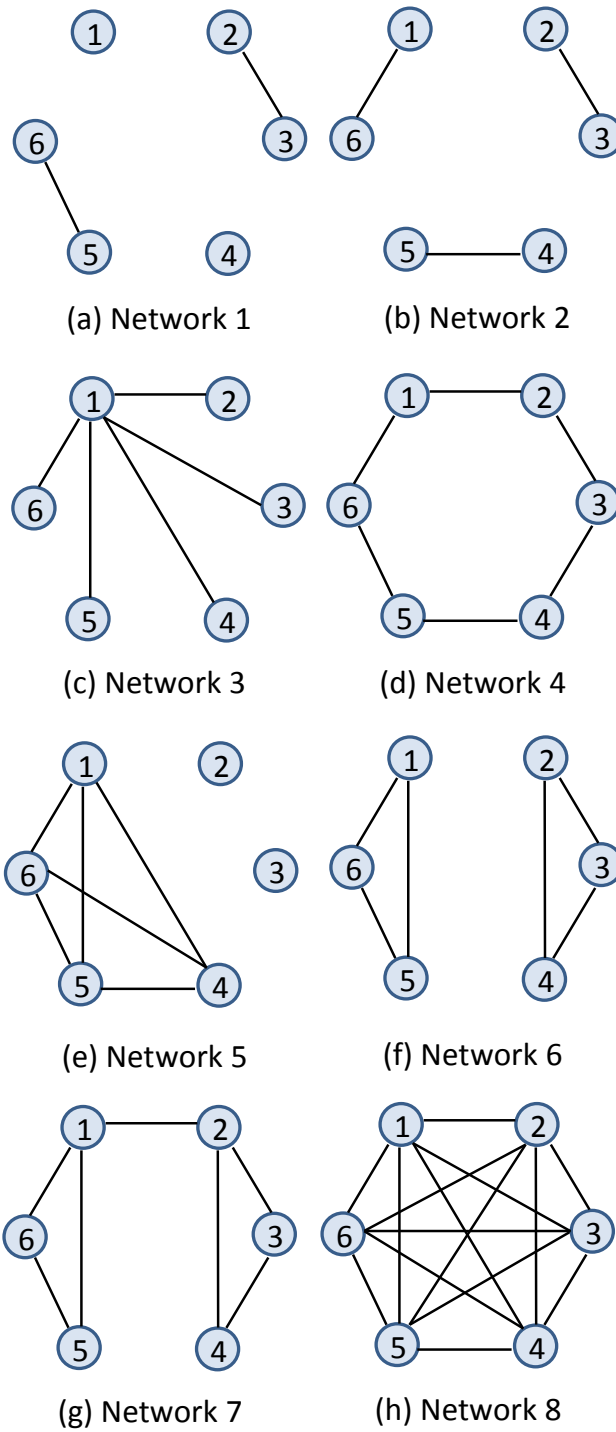


Figure 3.9: Eight Rössler networks.

In order to assess the effect of the number of oscillators in the computed synchrony values, we

constructed two subnetworks consisting of 3 oscillators each (as in Figure 3.9 (f)) and increased the number of oscillators in the network to 9 and 12. Table 2 shows the results for both HTS and S-estimator as the number of oscillators increases. Note that the first case, 6 oscillators, is the same as network 6 in Figure 3.9. For both methods, as the number of oscillators increases the multivariate synchrony decreases as there are more non-synchronized oscillators in the network. This trend is also observed from the unbiased estimators of HTS^2 and S^2 .

Finally, we assessed the effect of the number of subnetworks on the multivariate synchrony measures. Table 3 shows the results for HTS, S, their squared unbiased estimators for different number of subnetworks of three oscillators in a network of 12 oscillators. As expected, increasing the number of subnetworks increases the multivariate synchrony for both estimators.

Table 3.1: Multivariate synchrony (mean \pm st.dev.) in networks of Rössler oscillators.

Network	HTS	S	HTS_{UB}^2	S_{UB}^2
1	0.226 \pm 0.023	0.267 \pm 0.004	0.051 \pm 0.010	0.254 \pm 0.002
2	0.612 \pm 0.028	0.474 \pm 0.016	0.375 \pm 0.034	0.384 \pm 0.009
3	0.944 \pm 0.029	0.851 \pm 0.059	0.892 \pm 0.055	0.765 \pm 0.086
4	0.985 \pm 0.000	0.945 \pm 0.001	0.971 \pm 0.000	0.907 \pm 0.001
5	0.800 \pm 0.009	0.580 \pm 0.012	0.640 \pm 0.015	0.523 \pm 0.007
6	0.765 \pm 0.027	0.673 \pm 0.011	0.586 \pm 0.042	0.610 \pm 0.005
7	0.980 \pm 0.001	0.929 \pm 0.002	0.960 \pm 0.002	0.881 \pm 0.004
8	0.999 \pm 0.000	0.996 \pm 0.000	0.998 \pm 0.000	0.993 \pm 0.000

Table 3.2: Multivariate synchrony (mean \pm st.dev.) for different number of oscillators containing two subnetworks of three oscillators.

Number of Oscillators	HTS	S	HTS_{UB}^2	S_{UB}^2
6	0.765 \pm 0.027	0.673 \pm 0.011	0.585 \pm 0.042	0.610 \pm 0.005
9	0.682 \pm 0.017	0.457 \pm 0.011	0.463 \pm 0.024	0.372 \pm 0.009
12	0.613 \pm 0.015	0.348 \pm 0.008	0.373 \pm 0.018	0.263 \pm 0.006

Table 3.3: Multivariate synchrony (mean \pm st.dev.) for a network consisting of 12 oscillators for different number of subnetworks composed of three oscillators.

Number of Subnetworks	HTS	S	HTS_{UB}^2	S_{UB}^2
1	0.531 \pm 0.009	0.252 \pm 0.009	0.278 \pm 0.010	0.156 \pm 0.006
2	0.613 \pm 0.015	0.348 \pm 0.008	0.373 \pm 0.018	0.263 \pm 0.006
3	0.735 \pm 0.031	0.506 \pm 0.010	0.538 \pm 0.018	0.398 \pm 0.010
4	0.757 \pm 0.012	0.594 \pm 0.013	0.571 \pm 0.025	0.492 \pm 0.012

3.6.5 Assessment of topographical sensitivity

The topographical sensitivity of multivariate phase synchrony measures is evaluated on a network consisting of 58 chaotic non-identical Colpitts oscillators [36]. Colpitts oscillators have been employed in the assessment of topographical sensitivity since they are chaotic non-symmetrical oscillators that generate irregular sinusoidal signals similar to EEG [108]. Eq. (3.42) describes the dynamics of oscillator i . In this network, oscillators are coupled through $x_2^{(i)}$, and C^{ij} indicates the coupling from oscillator j to oscillator i ; $k = 0.5$, and g , Q and α are chosen randomly between

the intervals [4.006, 4.428], [1.342, 1.483], and [0.949, 0.999], respectively.

$$\begin{aligned}
\frac{dx_1^{(i)}}{dt} &= \frac{g}{Q(1-k)}[\alpha(e^{-x_2^{(i)}} - 1) + x_3^{(i)}] \\
\frac{dx_2^{(i)}}{dt} &= \frac{g}{Qk}[(1-\alpha)(e^{-x_2^{(i)}} - 1) + x_3^{(i)}] + \sum_{j=1}^{58} C^{ij}(x_2^{(j)} - x_2^{(i)}) \\
\frac{dx_3^{(i)}}{dt} &= -\frac{Qk(1-k)}{g}[(x_1^{(i)} - x_2^{(i)}) - \frac{1}{Q}x_3^{(i)}].
\end{aligned} \tag{3.42}$$

A network consisting of bidirectional connections among electrodes 9, 10, 11, 18, 19, 20, 27, 28, and 29 with the nonlinear dynamics described (3.42) was simulated for 100 repetitions. Here, coefficients C^{ij} are set to zero for all i and j except for $i, j = \{9, 10, 11, 18, 19, 20, 27, 28, 29\}$, corresponding to electrodes F1, FZ, F2, FC1, FCZ, FC2, C1, CZ, and C2 in Figure 3.10. For this network, it is expected that electrode FCz will show the highest degree of multivariate synchrony. This network is simulated via the Heun method with $\delta t = 0.04$ s. The signals are 100 seconds-long, the first 40 seconds are discarded to remove transients and signals are down-sampled to a sampling frequency of 12.5 Hz. In addition, white Gaussian noise was added and multivariate synchrony was assessed for SNR levels of 0 dB, 10 dB and 20 dB. S-estimator and HTS were computed at each electrode, for each time and frequency point among each electrodes first nearest neighbors, resulting in groups of five electrodes. For example, for this simulated network the S-estimator and HTS at electrode FCZ consider electrodes FZ, FC1, FCZ, FC2, and CZ.

Figure 3.11 shows the topographical plots for HTS and S-Estimator for the three SNR levels considered. Overall, HTS and S-estimator result in similar topographical maps, demonstrating good topographical sensitivity. Under additive noise, HTS results in higher multivariate synchrony among the electrodes comprising the simulated network, being more robust to noise when com-

pared to S-estimator.

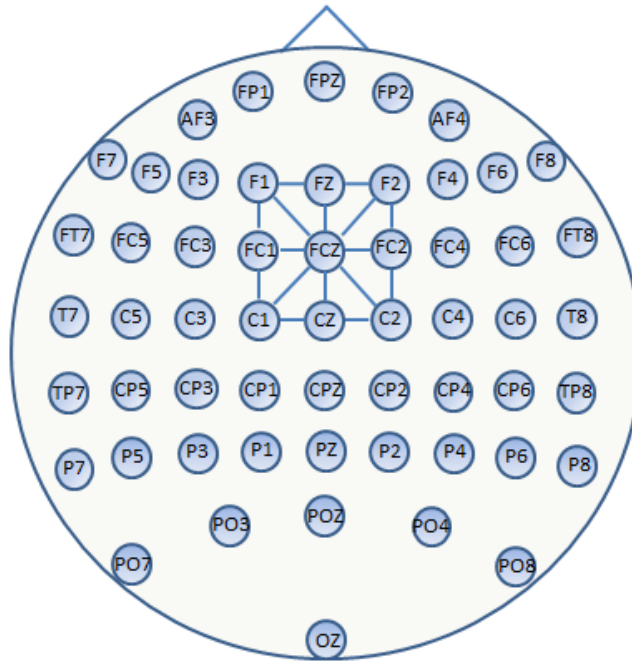


Figure 3.10: SynAmps2-64 EEG system and network under test.

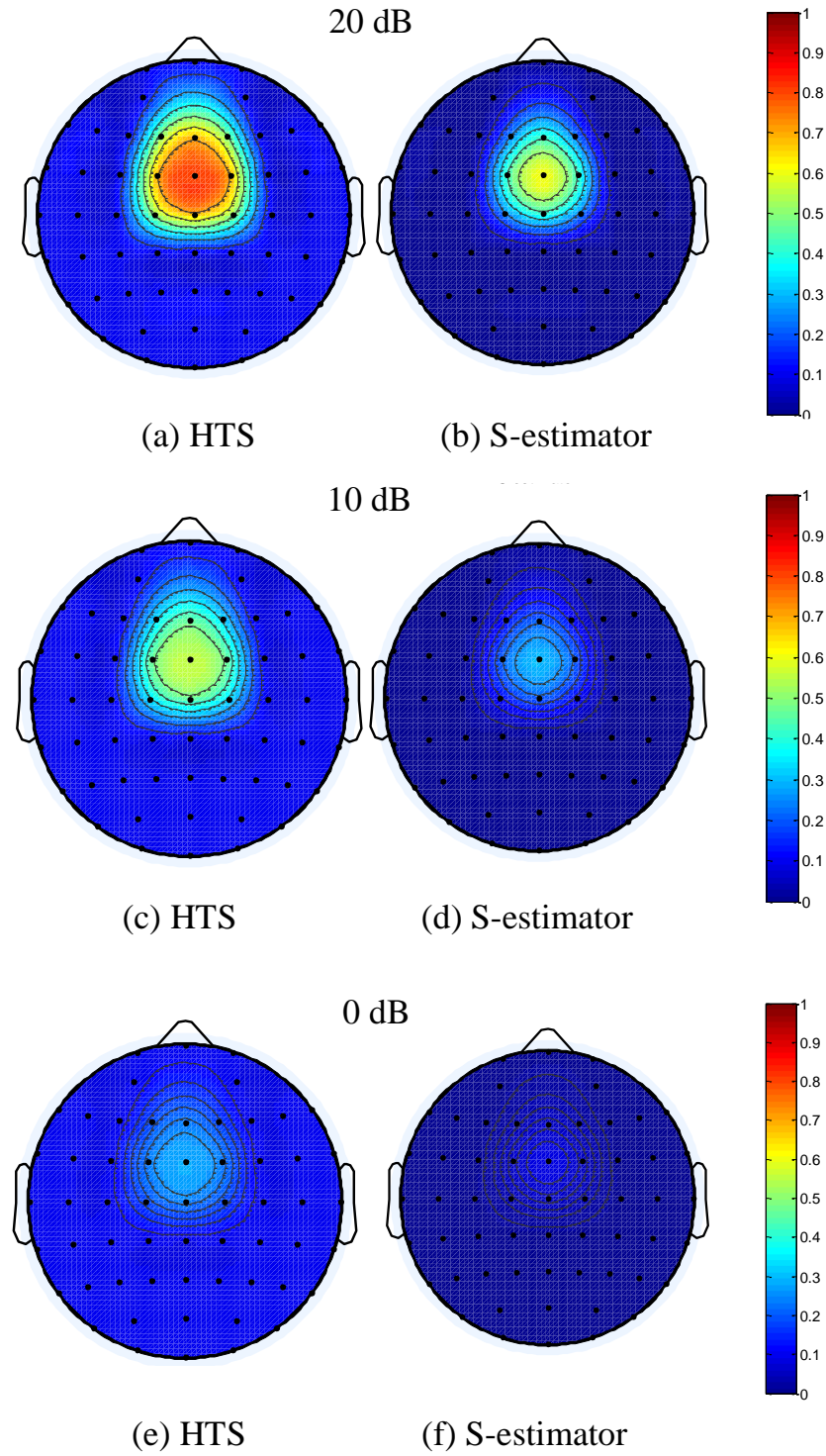


Figure 3.11: Topographical plots showing the sensitivity of HTS and S-estimator at various SNR levels. a), c) and e) HTS, 20 dB, 10 dB and 0 dB, respectively; b), d) and f) S-estimator, 20 dB, 10 dB and 0 dB, respectively.

First, we were interested in comparing the proposed multivariate measure with respect to conventional bivariate PLV. For this purpose, we compared the multivariate synchrony among electrodes FCz, F5 and F6 with the pairwise synchrony between FCz and F5 and FCz and F6 obtained from both error and correct responses for the 25-75 ms time interval, as this time interval contains the peak of the ERN and has been used in prior phase synchrony studies [29], [27]. These electrode pairs were selected as they are commonly used in assessing error-correct synchronization differences during cognitive control [27], [109]. The statistical significance of the difference between error and correct synchrony was investigated by performing a t-test for HTS and a two-way ANOVA for the two PLVs. Results are shown in Table 1 and indicate that HTS is more sensitive to the difference between error and correct responses across the central and lateral frontal regions when compared to PLV. This example shows the benefit of computing multivariate synchronization over bivariate pairs.

Table 3.4: Statistical significance of error-correct responses obtained from PLV and HTS.

Synchrony	p-value
PLV: FCz-F5	0.0209
PLV: FCz-F6	
HTS	1.019e-10

Topographical connectivity for EEG data is investigated following the same definition of electrode neighborhoods as in Section 3.6.5. HTS and S-estimator were computed for both error and correct responses separately, at each electrode as described in Section 3.6.5, obtaining a multivariate synchrony value for each time and frequency point and electrode for each subject. Next, for each subject, the multivariate synchrony was averaged over different time intervals of interest, and frequency bins corresponding to the theta band (4-8 Hz) at each electrode. Here, we inves-

tigated the dynamic nature of functional connectivity by looking at multivariate synchronization over different post-response time intervals. Figure 3.12 shows the topographical distribution of multivariate synchrony for error minus correct responses averaged over subjects estimated from HTS and S-estimator for four time intervals: 0-25 ms, 25-50 ms, 50-75 ms and 75-100 ms, and in the theta frequency band. From these figures, it is observed that HTS results in higher synchrony for error-correct difference for the frontal and central electrodes when compared to the central-parietal electrodes for all intervals. The S-estimator, on the other hand, does not indicate much variation across time and the error-correct synchrony differences are close to zero for most brain regions. For HTS, time intervals of 25-50 ms and 50-75 ms show moderate increase in synchrony in the central-frontal regions compared to the other time windows.

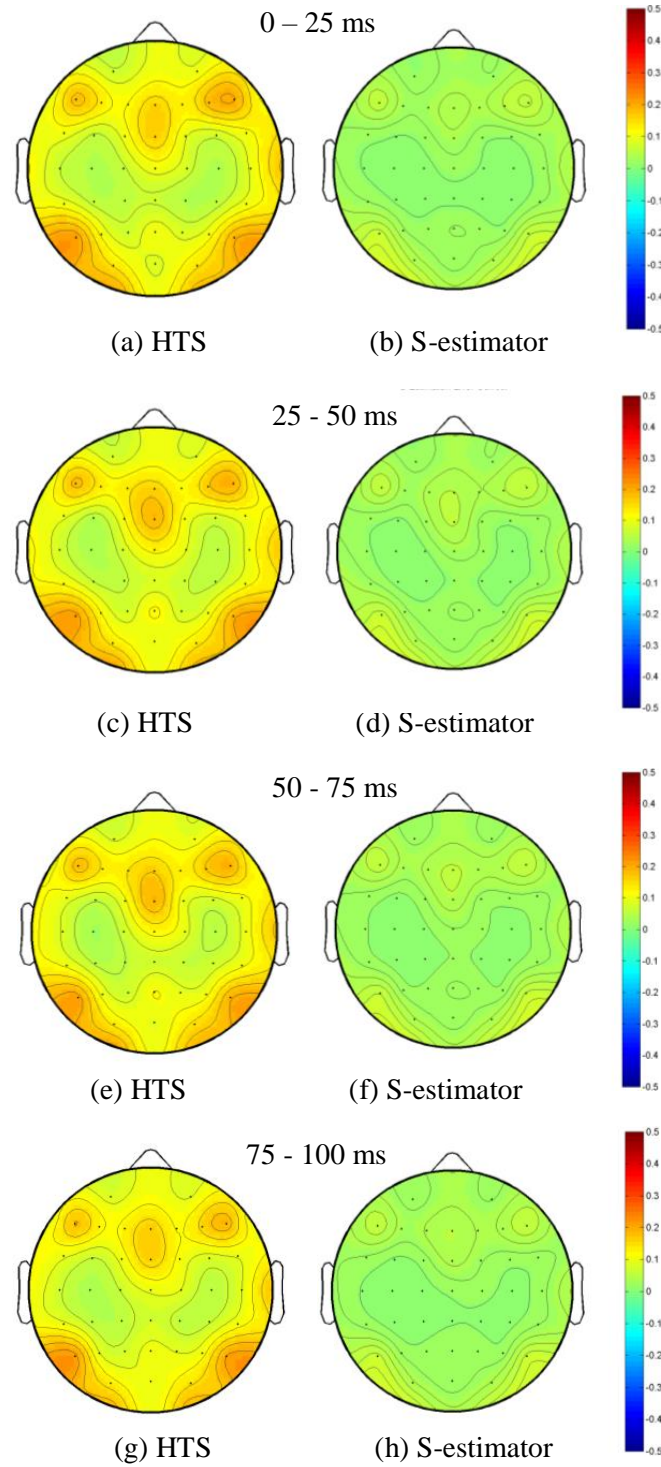


Figure 3.12: Error-Correct topographical sensitivity of multivariate synchrony in intervals of 25 ms and theta band [4-8 Hz]. (a), (c), (e) and (g) Error-Correct HTS difference. (b), (d), (f) and (h) Error-Correct S-estimator difference.

In order to investigate the significance of changes across time, t-tests were performed between consecutive time intervals. Figure 3.13 shows the topographical distribution of p-values obtained for each electrode for different comparisons. Clearly, there is a significant change in multivariate synchrony in the central electrodes between the first and second time intervals, i.e. 0-25 ms, and 25-50 ms (Figure 3.13 (a)), no significant change between 25-50 ms and 50-75 ms (Figure 3.13 (b)), and finally a small change in central electrodes between 50-75 ms and 75-100 ms (Figure 3.13 (c)).

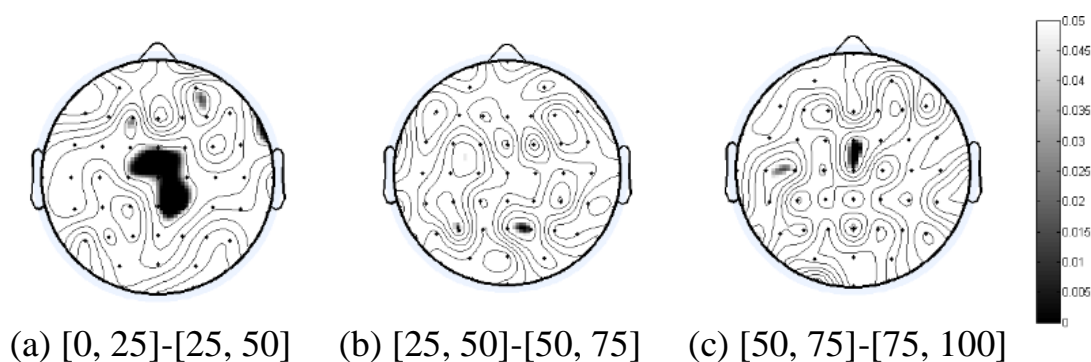


Figure 3.13: Topographical plots of p-values from t-test investigating the difference of multivariate synchrony from HTS between intervals. Note: Black regions correspond to lower p-values.

Since the intervals 25-50 ms and 50-75 ms show the largest error-correct differences based on Figure 3.12, we next focus on the time interval 25-75 ms and look at the topographical distribution of HTS for error and correct responses, separately. As Figure 3.14 shows, HTS yields increased synchronization for the central and lateral frontal regions for the error response (Figure 3.14 (a)) whereas there is no topographical differentiation of synchrony for the correct response (Figure 3.14 (b)). The proposed HTS measure replicates these findings and further identifies regionally increased synchronization on error trials, compared to correct, in the medial-lateral and frontal-central areas. Unlike conventional bivariate measures which require the computation of synchrony

nization between multiple pairwise regions, the proposed measure can summarize the global synchronization within a region by a single number providing an ease of interpretation.

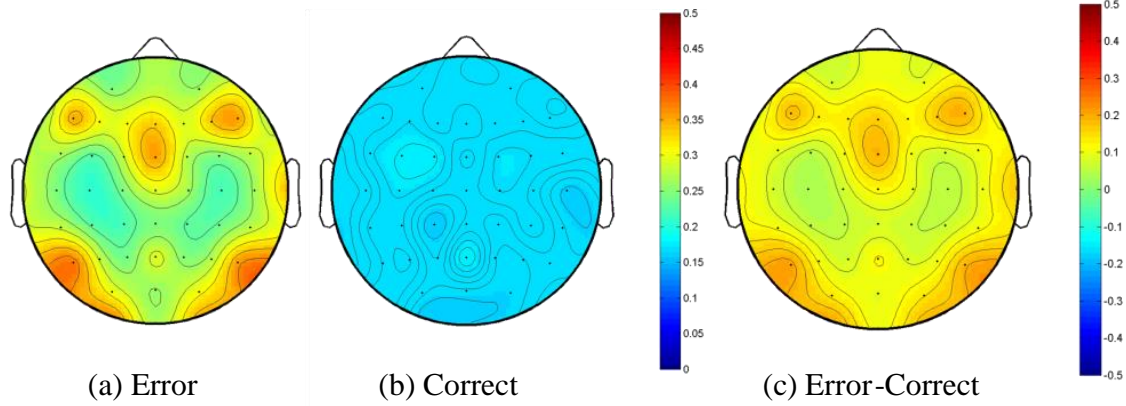


Figure 3.14: Topographical plots of multivariate synchrony from HTS in the 25-75 ms interval. (a) Error responses; (b) Correct responses; (c) Error-Correct responses.

In addition, a Receiver Operating Characteristic (ROC) curve was constructed in order to compare the performance of HTS and the S-estimator in the detection of multivariate synchrony during the ERN interval in the theta band. The probability of detection and false alarm were defined as the ratio at which the average multivariate synchrony over the ERN interval and theta band in electrodes FCz and CPz, respectively, exceeded a threshold. Figure 3.15 shows the ROC curves for both estimators. The area under the curve (AUC) for each estimator was computed, resulting in $AUC_{HTS} = 0.8601$ and $AUC_{S-estimator} = 0.7936$. Thus, as observed, HTS exceeds S-estimator in the detection of multivariate synchronization in the frontal-central regions during the ERN in the theta band indicating that HTS is more sensitive to detecting the difference in synchronization between the frontal-central region and the central-parietal region.

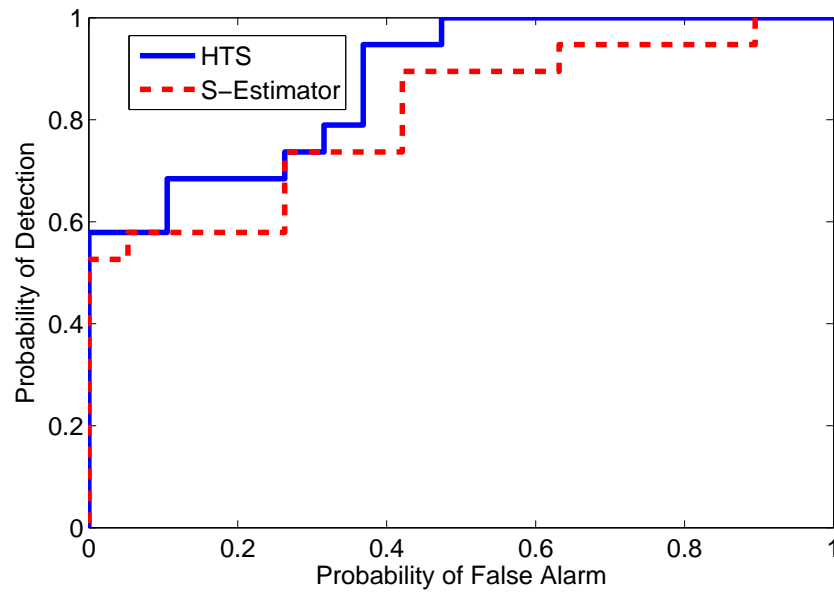


Figure 3.15: ROC curves for HTS and S-Estimator. Probability of detection is based on the multivariate synchrony among FCz and its neighbors whereas the probability of false alarm is based on the multivariate synchrony around CPz.

To further examine the functional, behavioral significance of the increased HTS synchrony identified following errors, compared to corrects, we computed correlations between the error-correct HTS synchrony and behavioral adjustments observed after mistakes post-error slowing (PES) calculated as correct response time (RT) following error responses correct RT following correct responses, and post-error accuracy (PEA) calculated as accuracy following error responses accuracy following correct responses. Although the functional significance of PES remains controversial i.e., some believe it represents the cautious slowing to ensure effective responding following mistakes, whereas others argue it represents off-task orienting to rare mistakes PEA is more clearly an adaptive response following errors (see [110] for a review). Figure 3.16 shows the topographical distribution of these correlations. Across frontal and central regions of interest, synchrony was inversely related to PES and positively related to PEA. The negative correlations between multivariate synchrony and PES were more broadly distributed whereas those between

multivariate synchrony and PEA were more localized to the central and left frontal electrodes. Figure 3.17 illustrates the topographical distribution of p-values (two-tailed) from the correlations obtained from PES (Figure 3.17 (a)) and PEA (Figure 3.17 (b)). For each behavioral condition, a distribution of correlations $p_{\Gamma}(\gamma)$ is constructed and a p-value for each electrode is obtained as $p - value_i = 2\min\{P(\Gamma \leq \gamma_i), P(\Gamma \geq \gamma_i)\}$ where γ_i denotes the correlation coefficient at electrode i . As shown in Figure 3.17 (a), the inverse relationship between HTS and PES is most significant in the right frontal (F6) and parietal regions (P6), whereas as shown in Figure 3.17 (b) correlation between HTS and PEA is most significant around the left -central (FC3, FC5, C3) and frontal regions (AF3).

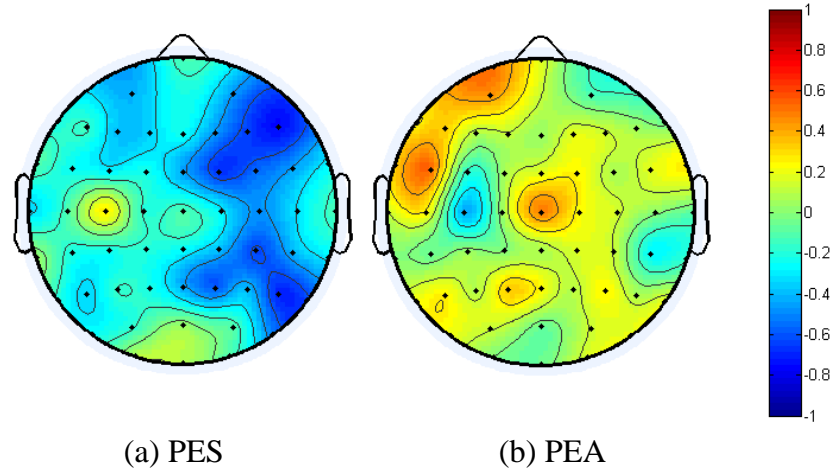


Figure 3.16: Correlation coefficient between (a) PES, (b) PEA and error-correct multivariate synchrony difference computed using HTS in the ERN interval 25-75 ms, for each electrode.

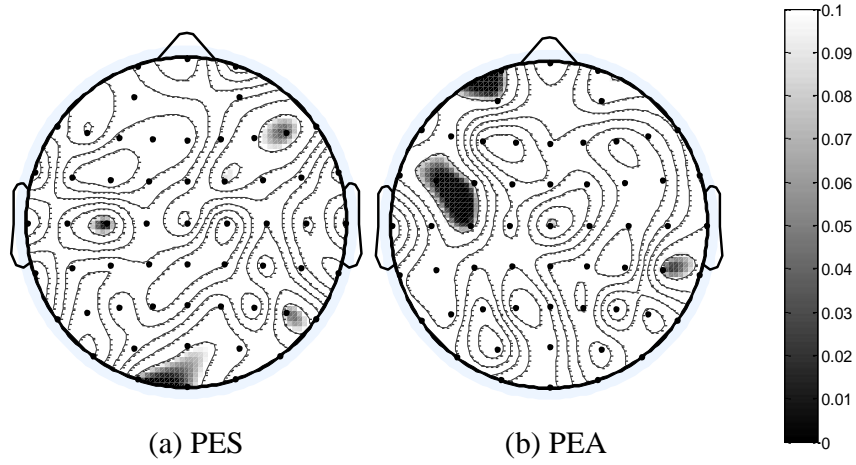


Figure 3.17: Topographical distribution of p-values obtained from the correlation coefficient between (a) PES, (b) PEA and error-correct multivariate synchrony difference computed using HTS in the ERN interval 25-75 ms, for each electrode. Black refers to more significant.

3.7 Conclusions

In this chapter, we presented a novel time-frequency measure of multivariate phase synchrony based on a hyperdimensional coordinate system. This measure has been derived from both a hyperspherical coordinate system and from the Cartesian product of unit circles. The proposed measure has been shown to be advantageous over a widely used multivariate measure, the S-estimator, in estimating the global synchrony in simulated systems of coupled oscillators and in neurophysiological signals. In particular, it was shown that the proposed method is a direct measure of global synchrony which overcomes the drawbacks of multivariate synchrony methods based on the bivariate PLV. First, it was shown, from a simulation in Rössler oscillators, that the proposed measure provides information about the underlying structure of the network, otherwise misinterpreted from the S-estimator. Second, the proposed measure is computationally efficient since it does not require the computation of all pairwise synchrony values in a network nor the eigendecomposition

of a connectivity matrix.

The application of this method to EEG data showed that it is more sensitive to the increase of multivariate synchrony among electrodes in the central-frontal region during error responses in an ERN experiment compared to the S-estimator. Furthermore, HTS was shown to be more sensitive to spatial changes in multivariate synchrony compared to S-estimator and to be a better predictor of error-correct differences in the error monitoring experiment compared to traditional bivariate PLV synchrony metrics. The proposed measure can be implemented using instantaneous phase estimates obtained from the Hilbert transform, the Wavelet transform and, with some limitations, the Hilbert-Huang transform in addition to the RID-Rihaczek distribution. Thus, the proposed measure of multivariate synchrony is a promising tool for the assessment of the global integration in dynamic complex networks.

Chapter 4

Graph to Signal Transform Based on the Resistance Distance and its applications to Functional Connectivity Networks

4.1 Introduction

Complex networks arise in a wide variety of systems such as biological, computational and social. In biology, for example, protein-protein interactions constitute protein interaction networks in which proteins represent nodes and their interactions are represented by edges [111]. Another example of complex networks is the Internet, for which nodes may represent computers, or webpages [112]. In the context of social networks, users are represented by nodes and their connections to other users are represented by edges [113]. Over the last decade, complex network theory has contributed significantly to the study of functional connectivity networks (FCNs), in particular in the assessment of functional integration and segregation [10]. Specifically, graph theoretic measures such as the shortest path length and clustering coefficient have helped to characterize small-world networks [114], and the degree distribution has been utilized to characterize scale-free networks

[115]. In particular, it has been shown that FCNs of the healthy population exhibit small-world properties [40], [116].

Despite the contributions of graph theoretic and information theoretic methods to the characterization of FCNs, these methods possess certain drawbacks. The first problem is the sensitivity of graph theoretic measures to network size. An example is the clustering coefficient, which quantifies the ratio of the number of triangles around a node and the maximum number of edges that can be connected to it [117]. Therefore the mean clustering coefficient can be unfairly affected by nodes with a low degree [117]. This would have an effect on other measures such as the small-world measure which depends on the clustering coefficient, particularly in brain networks constructed from electrophysiological modalities where the number of nodes tends to be small [118]. Another problem with graph theoretic measures is their non-uniqueness. An example is the small-world measure, which relies on the clustering coefficient normalized by the clustering coefficient of a random network. Such a normalization may affect the small-world measure as two very different network structures may have similar clustering coefficients [119]. Finally, another problem with graph theoretic measures is the mismatch between the measure and the flow of information in the underlying network, especially for weighted networks such as FCNs. For example, FCNs may not necessarily rely on shortest paths for communication between the nodes, and measures like the characteristic path length and the global efficiency are unable to capture this type of connectivity patterns [10], [120].

Alternatively, a complementary set of methods for network analysis has been proposed through transforming graphs into signals in order to take advantage of signal processing methods in the analysis [44], [43], [46]. By transforming graphs into signals it is possible to apply traditional signal processing techniques on those signals in order to extract information from the networks and overcome some of the shortcomings of graph measures. Both probabilistic and determinis-

tic methods have been proposed to transform networks into signals. In [121], a transformation of networks based on random walk theory has been proposed to show that the transformed signals reveal mixing patterns of the network. In another recent work, Girault et al. [122] proposed a semi-supervised learning method for graph to signal mapping which results in smooth signals from graphs. However, stochastic methods do not provide the means for recovering the network once they are transformed into signals. Shimada [44] and Haraguchi [43] formulated a deterministic method based on classical multidimensional scaling (CMDS), allowing the transformation of complex binary networks into time series. Under this transformation, the vertices of the network correspond to time indices for the time series [123]. It was shown that lattice networks transform to sinusoids and Watt-Strogatz networks transform to random signals [44]. Recently, Hamon et al. [124], [125] have extended this method to the analysis of temporal networks, with an application to a network of face-to-face contacts revealing significant subnetworks. However, all of these approaches have focused on binary graphs, and therefore have limited applicability to weighted networks that arise in neuroscience.

In order to transform both binary and weighted graphs into signals, we propose to use the resistance distance of a connected graph as the distance matrix for CMDS. The resistance distance of a graph was proposed in [126] in the context of chemistry. The resistance distance between two nodes corresponds to the equivalent resistance between them, considering the graph as an electric circuit [126] whose edges represent resistors inversely proportional to the edge weight. An advantage of the resistance distance over other graph distances, such as the shortest path distance, is that the resistance distance takes into account the global structure of the graph and hence reflects information about multiple paths. Moreover, the resistance distance can be obtained from the pseudoinverse of the Laplacian matrix of the graph and is a valid distance matrix [127]. Therefore, it is an alternative distance matrix for CMDS.

Transforming graphs into signals from CMDS results in a total of N signals or components, corresponding to each one of the N nodes in the network. We are interested in quantifying structural information of the networks based on their signals. In this work, we propose a graph entropy measure based on the spectrum of the signals obtained from the transformation. Graph information theoretic measures, such as graph entropy, are important as they allow for the quantification of the structural information of the networks. Quantifying the information content of FCNs through graph information measures has been limited. In one such study, Sato et al. [128] characterized FCNs from Attention Deficit Hyperactivity Disorder (ADHD) based on the graph spectral entropy of FCNs. In addition, Takahashi et al. [129] introduced the Jensen-Shannon divergence between graph spectra to compare networks from ADHD. Current graph entropy measures rely on the spectral distribution of the graph adjacency matrix or other graph-related matrices [130], and methods that consider the probability distribution on the graph vertices [131]. However, these methods present some drawbacks. In particular, spectral graph entropy may fail to discriminate the network structure among networks sharing similar spectra [132]. Moreover, graph entropy measures based on the probability distribution on the graph vertices rely heavily on the method for estimating the graph-vertex probability distributions. One method is based on partitioning the vertex set of binary graphs [133], [134]. Under this approach, vertices are clustered into sets of identical vertices based on their local and non-local degree-dependencies and a probability is assigned to each partition based on the total number of vertices in that partition divided by the total number of vertices [135]. However, this method relies on arbitrary parameters, which can be found in an optimal sense if there is a ground truth. Another method is based on an information functional for undirected and connected graphs [131]. In this method, the information functionals are function of the sets denoted j -sphere for each vertex, which is the set of vertices that are j edges apart from the current vertex where the distance is quantified by the shortest path. In this chapter, we propose

a graph entropy and divergence measure by implementing traditional information-theoretic measures on the spectrum of the signals obtained from the graph to signal transformation, and both methods are parameter free. It is shown that this method allows for the quantification of network structural information and discrimination between distinct cognitive network structures.

As the signals obtained from this graph transformation convey structural information of the networks, we consider them for event detection in temporal networks. Previously, Hamon et al. [124] used nonnegative matrix factorization of the spectra of the signals for characterizing the structure of temporal networks. By doing this, it is possible to characterize the network structure over time, which may remain ambiguous if assessed from graph theoretic measures solely. Another important problem in the study of temporal networks is the detection of events, which occur as a deviation of the network structure from its usual structure at a particular instance. Due to the evolving nature of temporal networks in many applications, detecting abrupt changes in the structure of temporal networks is of great importance. In particular, the connections among the nodes of the network may change with time or some network components may be removed over time [136]. Previously proposed methods for event detection in temporal networks include [136] distance based methods, probabilistic model based methods, and subspace estimation-based methods. In distance based methods, the distance between two graphs is computed at each time point and thus anomalous instances are extracted from this time series. Probabilistic model based methods account for the deviations from models of the graph spectrum, which indicate the presence of an event. Subspace estimation based methods, such as singular value decomposition and tensor decompositions, track the singular values over time as well as the reconstruction error for identifying events in the temporal networks. On the other hand, several methods have been proposed in computer science and sensor networks literature, with applications to video and wireless sensors. In addition, work on the analysis of temporal networks has focused on the computation of graph theoretic measures

based on the graph adjacency matrix at each time point and then constructing time series for each feature [137]. However, it has been argued that this approach may fail when activity patterns are discontinuous and when there are abrupt changes in the network [138].

In this work, we focus on tensor decompositions since they exploit the underlying relationships of multiway data [139], as opposed to matrix decomposition methods, such as PCA and Independent Component Analysis (ICA). In this chapter, we propose to form a tensor based on the spectra of the signals obtained from the graph to signal transformation at each time point and detect abrupt structural changes in the temporal network at a particular time. Our proposed method is compared to the tensor decomposition based on the graph adjacency matrices over time. Both methods correctly detect the change points in time when significant changes occur, however, the proposed method is more sensitive to those changes.

Finally, we employ the proposed graph to signal transformation for characterizing functional connectivity network structure. The proposed measure is applied to the electroencephalogram (EEG) data described in Chapter 2. Previous works have suggested that functional connectivity networks are small-world networks [140]. Furthermore, a recent work on weighted small-world measures have shown increased small-world characteristics in functional connectivity networks during error-related negativity [116]. In this chapter, we assess the small-worldness of functional connectivity networks during ERN by computing the correlation between the spectrum of the signals transformed and the spectrum of a small-world network for different parameters of small-worldness. We show that the signals obtained from graphs contain structural information that allows us to demonstrate the structural differences between different experimental conditions. In particular, we show how the structure of functional connectivity networks from different conditions in a cognitive control experiment can be characterized with the proposed method.

This chapter is organized as follows. Section 4.2 presents a background on graph entropy

measures and CMDS for graph to signal transform of binary networks. Next, the proposed graph to signal transform based on the resistance distance matrix is presented in Section 4.3 as well as characterizations of this transform. Simulation results comparing the proposed method to the previously defined distance measures are shown in Section 4.3.5.1. In 4.4 Next, we present the proposed graph entropy measures based on the signals obtained from the transform in Section 4.5, followed by event detection in Section 4.6 and the characterization of functional connectivity networks in Section 4.7. Finally, conclusions are discussed in Section 4.8.

4.2 Background

4.2.1 Graph Entropy Measures

Previously proposed graph entropy measures to quantify the structural information of the graph include information functionals evaluated at the local node level [131], and the spectrum of the adjacency matrix \mathbf{A} or the Laplacian [130]. Let $d(u, v)$ be the shortest path between nodes u and v and $\sigma(v) = \max_{u \in V} d(u, v)$ be the eccentricity of the graph. Define $\rho(G) = \max_{v \in V} \sigma(v)$ as the diameter of the graph and $S_j(v_i) = \{v \in V | d(v_i, v) = j\}$, $j \geq 1$, as the j -sphere of node v_i . In this work, we compare our proposed method to two information functionals.

The first information functional is defined as $f^{V_1}(v_i) = \alpha^{c_1|S_1(v_i)| + c_2|S_2(v_i)| + \dots + c_{\rho(G)}|S_{\rho(G)}(v_i)|}$, where $c_k > 0$, $1 \leq k \leq \rho(G)$, $\alpha > 0$, and V_1 identifies it as the first information functional. This functional is a function of the j -sphere for each node. Node probabilities are defined based on this functional as $p^{V_1}(v_i) = \frac{f^{V_1}(v_i)}{\sum_{j=1}^N f^{V_1}(v_j)}$. Based on p^{V_1} , the graph entropy of G is given as

$$I_{f^{V_1}}(G) = - \sum_{i=1}^{|V|} p^{V_1}(v_i) \log(p^{V_1}(v_i)). \quad (4.1)$$

The second graph entropy measure that is commonly used is obtained from the spectrum of the adjacency matrix \mathbf{A} [130]. In this case, the information functional is defined as $f_i^{V_2} = |\lambda_i|$, where λ_i is the i th eigenvalue of the adjacency matrix. The node probability is defined as $p_i^{V_2} = \frac{|\lambda_i|}{\sum_{j=1}^C |\lambda_j|}$, where $\lambda_1, \dots, \lambda_C$ correspond to the non-zero eigenvalues of the adjacency matrix. Then, the graph entropy is defined as

$$I_{f^{V_2}}(G) = - \sum_{i=1}^C p_i^{V_2} \log(p_i^{V_2}). \quad (4.2)$$

This functional is equivalent to the Von Neumann entropy of a graph if we consider the eigenvalues of the normalized Laplacian [141], [142].

Finally, the one-dimensional structural information of a connected, undirected graph has been defined based on the degree distribution as [143]

$$\mathcal{H}^1(G) = - \sum_{i=1}^N \frac{\Delta_{ii}}{2M} \log \frac{\Delta_{ii}}{2M}, \quad (4.3)$$

where Δ_{ii} corresponds to the degree of the i th node, N is the total number of nodes in the network, and M is the total number of edges in the network.

4.2.2 Graph to Signal Transform based on Classical Multidimensional Scaling

The transformation of graphs into signals is based on CMDS. CMDS is a data reduction algorithm whose objective is to find a low-dimensional representation of the data while preserving the Euclidean distances between points [144]. In particular, for our application of transforming graphs into signals, the aim is to obtain coordinate vectors that preserve a given distance [43]. The first

step in the algorithm is to entry-wise square the distance matrix \mathbf{D} and double center it as

$$\mathbf{B} = -\frac{1}{2}\mathbf{J}_N\mathbf{D}^{(2)}\mathbf{J}_N, \quad (4.4)$$

where $\mathbf{D}^{(2)} = \mathbf{D} \circ \mathbf{D}$ is the entry-wise squared Euclidean distance matrix, $\mathbf{J}_N = \mathbf{I}_N - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N'$ is a centering matrix, \mathbf{I}_N is an $N \times N$ identity matrix, $\mathbf{1}_N$ is a $N \times 1$ vector of ones, and $'$ denotes the transpose. \mathbf{B} is a positive semidefinite matrix with $\text{rank}(\mathbf{B}) = C$, $C \leq N$. Therefore, \mathbf{B} has C positive eigenvalues, and $N - C$ eigenvalues equal to zero. The next step is to perform the spectral factorization of \mathbf{B} , resulting in $\mathbf{B} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}' = \left(\mathbf{P}\mathbf{\Lambda}^{\frac{1}{2}}\right) \times \left(\mathbf{P}\mathbf{\Lambda}^{\frac{1}{2}}\right)' = \mathbf{X}\mathbf{X}'$, where $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_C)$, and $\mathbf{\Lambda}^{\frac{1}{2}} = \text{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_C})$, correspond to the nonzero eigenvalues of \mathbf{B} , with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_C$, $\mathbf{P} \in \mathbb{R}^{N \times C}$, and $\mathbf{X} \in \mathbb{R}^{N \times C}$. Based on \mathbf{X} , a total of C signals of length N corresponding to the columns of \mathbf{X} are obtained. The i th signal $\mathbf{x}_i \in \mathbb{R}^{N \times 1}$ is defined as the i th column of \mathbf{X} with $i = 1, 2, \dots, C$. In this chapter, we will refer to $\mathbf{x}_i(n)$ as components and signals interchangeably.

In order to preserve the positive definiteness of \mathbf{B} , the matrix \mathbf{D} needs to be a valid distance matrix and conditionally negative definite. In previous works [44], [124], CMDS has been implemented in the transformation of binary networks and the distance \mathbf{D} is based on the binary adjacency matrix \mathbf{A} as

$$\mathbf{D}_{ij} = \begin{cases} 0, & i = j, \\ 1, & a_{ij} = 1 \text{ and } i \neq j, \\ \gamma, & a_{ij} = 0 \text{ and } i \neq j, \end{cases} \quad (4.5)$$

where γ is a parameter that guarantees the conditional negative definiteness of \mathbf{D} . In this manner, the coordinate vectors preserve the adjacency relationship of the nodes [43]. In [46], Hamon et. al

found the upper bound of γ to be $\sqrt{\frac{N}{N-2}}$. It is important to note that \mathbf{D} only provides information about whether the vertices are connected or not, but not about the distance between vertices.

4.3 Graph to Signal Transformation Based on the Resistance Distance Matrix

4.3.1 Resistance distance

In order to extend the graph to signal transform based on CMDS to weighted graphs, we consider the resistance distance matrix of a graph, denoted as \mathbf{R} . The resistance distance was introduced by Klein and Randic [126] as an alternative to the shortest path distance for applications in chemistry. It is inspired by basic circuit theory, where each edge on the graph represents a resistor with value $\frac{1}{w_{ij}}$ [145]. The following definitions of the resistance distance apply to both binary and weighted networks. For simplicity, the notation will follow that of binary networks.

The resistance distance matrix $\mathbf{R} \in \mathbb{R}^{N \times N}$ of a complete graph with N nodes is given by

$$\mathbf{R} = \tilde{\mathbf{Z}}\mathbf{1}_N\mathbf{1}_N' + \mathbf{1}_N\mathbf{1}_N'\tilde{\mathbf{Z}} - 2\mathbf{Z}, \quad (4.6)$$

where $\mathbf{Z} = (\mathbf{L} + \frac{1}{N}\mathbf{1}_N\mathbf{1}_N')^{-1}$ and $\tilde{\mathbf{Z}} = \text{diag}(Z_{11}, Z_{22}, \dots, Z_{NN})$. In addition, \mathbf{R} is a conditionally negative matrix and therefore has only one positive eigenvalue.

Each entry R_{ij} in \mathbf{R} corresponds to the squared Euclidean distance between vertices i and j [127]. In a connected graph, $R_{ij} \leq d(i, j)$, where $d(i, j)$ is the shortest path distance, and equality holds when there is only one path between i and j [146]. The resistance distance is a valid distance measure, and satisfies the following properties [147]:

$$\begin{aligned}
R_{ij} &\geq 0 \text{ for all } i, j \text{ with equality if and only if } i = j, \\
R_{ij} &= R_{ji}, \\
R_{ij} + R_{jk} &\geq R_{ik}.
\end{aligned} \tag{4.7}$$

The resistance distance between vertices i and j , R_{ij} , is defined through the Moore-Penrose pseudo inverse of \mathbf{L} , \mathbf{L}^\dagger [44], as

$$R_{ij} = L_{ii}^\dagger + L_{jj}^\dagger - 2L_{ij}^\dagger. \tag{4.8}$$

Equivalently, R_{ij} can be computed as

$$R_{ij} = (\mathbf{e}_i - \mathbf{e}_j)^T \mathbf{L}^\dagger (\mathbf{e}_i - \mathbf{e}_j), \tag{4.9}$$

where \mathbf{e}_i and \mathbf{e}_j are $N \times 1$ vectors of zeros with 1 in the i th and j th index, respectively. In addition, the resistance distance is related to random walks on the graph, where the R_{ij} is proportional to the expected commute time of a random walk on the graph [146] and is given by

$$R_{ij} = \frac{1}{2|E|} [E(T^{ij}) + E(T^{ji})], \tag{4.10}$$

where T^{ij} is the number of transitions from vertex i to vertex j and $|E|$ is the cardinality of the edge set.

The resistance distance also provides a measure of network robustness. In particular, the effective resistance, also known as the Kirchhoff index, provides information regarding to how robust the

network is [146], [148], with a small effective resistance value indicating a robust network. The effective resistance distance is defined as

$$R_{eff} = \sum_{1 \leq i \leq j \leq N} R_{ij} = N \sum_{k=2}^N \frac{1}{\lambda_k} = N \text{tr}(\mathbf{L}^\dagger), \quad (4.11)$$

where $\lambda_i, i = 2, \dots, N$ correspond to the eigenvalues of \mathbf{L} . Finally, R_{eff} is related to the network criticality γ , $\gamma = 2R_{eff}$ [146], [149].

4.3.2 Classical Multidimensional Scaling based on the Resistance Distance

Using \mathbf{R} , (4.4) can be reexpressed as

$$\mathbf{B} = -\frac{1}{2} \mathbf{J}_N \mathbf{R} \mathbf{J}_N, \quad (4.12)$$

to obtain the signals \mathbf{X} for both binary and weighted networks. Note that since the entries of \mathbf{R} correspond to squared Euclidean distances, it is not necessary to square the entries of \mathbf{R} for the computation of \mathbf{B} .

As described in [43], if we denote the i th point in an m -dimensional Euclidean space as $\tilde{\mathbf{x}}_i = (\tilde{x}_{i1}, \tilde{x}_{i2}, \dots, \tilde{x}_{im})', i = 1, \dots, N$ and define the coordinate matrix as $\tilde{\mathbf{X}} = (\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \dots, \tilde{\mathbf{x}}_N)'$, then the Euclidean distance matrix in (4.12) is given as $\mathbf{R} = \text{diag}(\tilde{\mathbf{X}} \tilde{\mathbf{X}}' \mathbf{1}_N \mathbf{1}_N') + \mathbf{1}_N \mathbf{1}_N' \text{diag}(\tilde{\mathbf{X}} \tilde{\mathbf{X}}') - 2\tilde{\mathbf{X}} \tilde{\mathbf{X}}'$. Then, (4.12) can be reexpressed as

$$B = -\frac{1}{2}\mathbf{J}_N \mathbf{R} \mathbf{J}_N \quad (4.13)$$

$$\begin{aligned} &= -\frac{1}{2}\mathbf{J}_N [\text{diag}(\tilde{\mathbf{X}}\tilde{\mathbf{X}}' \mathbf{1}_N \mathbf{1}_N') + \mathbf{1}_N \mathbf{1}_N' \text{diag}(\tilde{\mathbf{X}}\tilde{\mathbf{X}}') - 2\mathbf{X}\mathbf{X}'] \mathbf{J}_N \\ &= \mathbf{J}_N \tilde{\mathbf{X}}\tilde{\mathbf{X}}' \mathbf{J}_N \\ &= \mathbf{X}\mathbf{X}', \end{aligned} \quad (4.14)$$

where $\mathbf{X} = \mathbf{J}\tilde{\mathbf{X}}$ is the matrix whose columns correspond to the signals from the network, as described in Section 4.2.2.

4.3.3 Reconstruction of the original graph

If the signals \mathbf{X} are not distorted, then in principle the resistance distance matrix \mathbf{R} can be recovered from the signals through the computation of the squared Euclidean distance between the points

$$\hat{R}_{ij} = \sum_{c=1}^C (x_c(i) - x_c(j))^2, \quad (4.15)$$

where $\hat{\mathbf{R}}$ is the estimated \mathbf{R} , C corresponds to the total number of components and $x_c(i)$ and $x_c(j)$ correspond to the i th and j th entries of the c th component. It is possible to recover the original adjacency matrix from $\hat{\mathbf{R}}$, for both weighted and binary graphs as follows. First, we introduce τ as [147]

$$\tau_i = 2 - \sum_{j \in V(i)} \frac{\hat{R}_{i,j}}{W_{ij}}, \quad (4.16)$$

where $V(i)$ denotes the set of vertices adjacent to i where adjacency is defined as $W_{ij} \neq 0$.

For the next step, since \mathbf{R} is nonsingular [127], we consider the following expression of $\hat{\mathbf{R}}^{-1}$

which follows from the inverse of Euclidean distance matrices [150]:

$$\hat{\mathbf{R}}^{-1} = -\frac{1}{2}\hat{\mathbf{L}} + \frac{1}{\tau'\hat{\mathbf{R}}\tau}\tau\tau'. \quad (4.17)$$

From (4.17), the Laplacian matrix $\hat{\mathbf{L}}$ is estimated as

$$\hat{\mathbf{L}} = -2\left(\hat{\mathbf{R}}^{-1} - \frac{1}{\tau'\hat{\mathbf{R}}\tau}\tau\tau'\right). \quad (4.18)$$

Given an estimate of the Laplacian matrix, the degree matrix $\hat{\Delta}$ is estimated as the diagonal matrix whose elements are the diagonal entries of $\hat{\mathbf{L}}$

$$\hat{\Delta} = \text{diag}(\hat{L}_{11}, \hat{L}_{22}, \dots, \hat{L}_{NN}). \quad (4.19)$$

Finally, the weighted adjacency matrix $\hat{\mathbf{W}}$ is computed as

$$\hat{\mathbf{W}} = \hat{\Delta} - \hat{\mathbf{L}}. \quad (4.20)$$

An alternative procedure for network reconstruction is as follows. It has been previously shown that \mathbf{R} is an Euclidean Distance matrix if and only if $\mathbf{B} = -\frac{1}{2}\mathbf{J}\mathbf{R}\mathbf{J}$ is a positive definite matrix [150]. If \mathbf{R} is an invertible Euclidean Distance matrix, then $\mathbf{R}^{-1} = -\mathbf{Y} + \mathbf{u}\mathbf{u}^T$, where \mathbf{Y} is a positive definite matrix, $\text{rank}(\mathbf{Y}) = N - 1$, and $\mathbf{1}^T\mathbf{Y} = 0$, that is, the sum of its rows is zero. As shown in [150], \mathbf{Y} corresponds to the Moore-Penrose pseudoinverse of \mathbf{B} . Therefore, $\hat{\mathbf{L}}^\dagger$ can be computed from $\hat{\mathbf{R}}$ as

$$\hat{\mathbf{L}}^\dagger = -\frac{1}{2}\mathbf{J}\hat{\mathbf{R}}\mathbf{J}. \quad (4.21)$$

Finally, $\hat{\Delta}$ and $\hat{\mathbf{W}}$ are estimated as before.

4.3.4 Perturbation Analysis

In this section, we describe analytically the effect of small perturbations such as addition and removal of edges to the network on the transformed signals. Specifically, we are interested in describing how (4.12) is affected in both cases and its effect on the graph to signal transformation. We begin with the case edges are added to the network, where the perturbed adjacency matrix is given as

$$\tilde{\mathbf{W}} = \mathbf{W}_o + \mathbf{W}_\delta, \quad (4.22)$$

where \mathbf{W}_o is the original adjacency matrix, \mathbf{W}_δ is an adjacency matrix with the new edges and $\tilde{\mathbf{W}}$ is the perturbed adjacency matrix. The following results apply to both weighted and binary graphs. We will follow the notation of weighted graphs for simplicity. The Laplacian of the perturbed network can be expressed in terms of the original network and the new links. As described by [149], the graph Laplacian can be reexpressed in terms of the graph incidence matrix as

$$\mathbf{L} = \mathbf{C}\mathbf{W}^E\mathbf{C}^T = \sum_{i,j} w_{ij}^E \mathbf{e}_{ij} \mathbf{e}_{ij}^T, \quad (4.23)$$

where $\mathbf{C} \in \mathbb{R}^{N \times M}$ is the graph incidence matrix, \mathbf{e}_{ij} is a column vector whose i th entry equals 1 and the j th entry equals -1, $\mathbf{W}^E \in \mathbb{R}^{M \times M}$ is a diagonal matrix whose entries are the M graph edges, and M is the total number of edges in the network. When edges are added to the network, (4.23) results in

$$\begin{aligned}
\tilde{\mathbf{L}} &= \tilde{\mathbf{C}}\tilde{\mathbf{W}}^E\tilde{\mathbf{C}}^T \\
&= \tilde{\mathbf{C}}_o\mathbf{W}_o^E\tilde{\mathbf{C}}_o^T + \tilde{\mathbf{C}}_\delta\mathbf{W}_\delta^E\tilde{\mathbf{C}}_\delta^T \\
&= \mathbf{L}_o + \mathbf{L}_\delta
\end{aligned} \tag{4.24}$$

where \mathbf{L}_o is the Laplacian matrix of the original network and \mathbf{L}_δ is the Laplacian matrix of the subnetwork corresponding to the perturbation. The Moore-Penrose pseudoinverse of (4.24) was found [149] to be

$$\tilde{\mathbf{L}}^\dagger = \mathbf{L}_o^\dagger - \mathbf{L}_o^\dagger \mathbf{L}_\delta (\mathbf{I} + \mathbf{L}_o^\dagger \mathbf{L}_\delta)^{-1} \mathbf{L}_o^\dagger. \tag{4.25}$$

This expression is valid as long as the addition of edges does not increase the rank of $\tilde{\mathbf{L}}$ more than the rank of \mathbf{L}_o . A similar expression is obtained for the case when edges are removed from the network. In this case, the perturbed adjacency matrix is expressed as

$$\tilde{\mathbf{W}} = \mathbf{W}_o - \mathbf{W}_\delta, \tag{4.26}$$

where now \mathbf{W}_δ is a matrix whose entries $W_{\delta ij}$ are identical to the edges e_{ij} to be removed from \mathbf{W}_o and zero otherwise. Thus, the Moore-Penrose pseudoinverse of the Laplacian can now be expressed as

$$\tilde{\mathbf{L}}^\dagger = \mathbf{L}_o^\dagger + \mathbf{L}_o^\dagger \mathbf{L}_\delta (\mathbf{I} + \mathbf{L}_o^\dagger \mathbf{L}_\delta)^{-1} \mathbf{L}_o^\dagger, \tag{4.27}$$

where the only difference between (4.25) and (4.27) is the sign and the same conditions on the

rank of \mathbf{L} hold.

Based on (4.25) and (4.27) we can express (4.12) for the case when edges are added or removed, respectively. We show only the derivation for the former scenario, since the only variant in the case of missing edges is a sign. We begin by recalling that $R_{ij} = L_{ii}^\dagger + L_{jj}^\dagger - 2L_{ij}^\dagger$. Denote \tilde{R}_{ij} the resistance distance between edges i and j for the perturbed network. Based on (4.25), we can express \tilde{R}_{ij} when edges are added as

$$\tilde{R}_{ij} = [\mathbf{L}_o^\dagger + \mathbf{L}_o^\dagger \mathbf{L}_\delta (\mathbf{I} + \mathbf{L}_o^\dagger \mathbf{L}_\delta)^{-1} \mathbf{L}_o^\dagger]_{ii} + [\mathbf{L}_o^\dagger + \mathbf{L}_o^\dagger \mathbf{L}_\delta (\mathbf{I} + \mathbf{L}_o^\dagger \mathbf{L}_\delta)^{-1} \mathbf{L}_o^\dagger]_{jj} - 2[\mathbf{L}_o^\dagger + \mathbf{L}_o^\dagger \mathbf{L}_\delta (\mathbf{I} + \mathbf{L}_o^\dagger \mathbf{L}_\delta)^{-1} \mathbf{L}_o^\dagger]_{ij}, \quad (4.28)$$

where $[\cdot]_{ij}$ denotes the i, j entry of the matrix $\tilde{\mathbf{L}}$. Let $\mathbf{M} = \mathbf{L}_o^\dagger \mathbf{L}_\delta (\mathbf{I} + \mathbf{L}_o^\dagger \mathbf{L}_\delta)^{-1} \mathbf{L}_o^\dagger$. Rearranging (4.28) we obtain

$$\begin{aligned} \tilde{R}_{ij} &= L_{1ii}^\dagger + L_{1jj}^\dagger - 2L_{1ij}^\dagger - (M_{ii} + M_{jj} - 2M_{ij}), \\ &= R_{ij}^o - R_{ij}^M, \end{aligned}$$

where R_{ij}^o refers to the resistance distance of the original network and R_{ij}^M is the resistance distance defined from the perturbation matrix \mathbf{M} . Now, we can define an alternative matrix $\tilde{\mathbf{B}}$ as

$$\begin{aligned}
\tilde{\mathbf{B}} &= -\frac{1}{2}\mathbf{J}_N\tilde{\mathbf{R}}\mathbf{J}_N \\
&= -\frac{1}{2}\mathbf{J}_N[\mathbf{R}^o - \mathbf{R}^M]\mathbf{J}_N \\
&= -\frac{1}{2}\mathbf{J}_N\mathbf{R}^o\mathbf{J}_N + \frac{1}{2}\mathbf{J}_N\mathbf{R}^M\mathbf{J}_N \\
&= \mathbf{B}_o - \mathbf{B}_\delta.
\end{aligned} \tag{4.29}$$

Therefore, we can express the matrix \mathbf{B} for CMDS based on the resistance distance matrix when edges are added/removed to the network in terms of the original network and the new set of links. The previous procedure can be similarly carried out to obtain an expression for the case when edges are removed as $\tilde{\mathbf{B}} = \mathbf{B}_o + \mathbf{B}_\delta$.

4.3.5 Illustration of Graph to Signal Transform

In this section, we first compare the graph to signal transformation based on the distance \mathbf{D} and \mathbf{R} for binary networks. Next, the reconstruction of weighted networks is assessed. Following this, the robustness of the proposed method is assessed for various types of network anomalies.

4.3.5.1 Binary graphs

Signals from binary graphs based on the resistance distance matrix are first compared to those obtained from CMDS based on the distance matrix \mathbf{D} . First, we simulate two ring networks with $N = 128$ nodes and average degree $K = 2$ and $K = 10$. Figure 4.1 shows the results obtained from both methods. As expected, the signals based on the resistance distance matrix are sinusoidal signals (Figure 4.1 (a) and Figure 4.1 (b)) similar to the signals obtained from \mathbf{D} (Figure 4.1 (c) and Figure 4.1 (d)). Figure 4.1 (a) and Figure 4.1 (c) show the first three components obtained

from \mathbf{R} and \mathbf{D} , respectively, for an average degree K equal to 2, and similarly Figure 4.1 (b) and Figure 4.1 (d) for $K = 10$. From these figures, it is observed that the amplitude of signals obtained from \mathbf{R} is inversely proportional to the average degree, K , with the maximum amplitude when $K = 2$ and reduced amplitude when $K = 10$. This can be explained in terms of the resistance distance. It has been shown previously that the pairwise resistance R_{ij} decreases when edges are added or weights are increased in the network [146]. In addition, it has been shown that the signals obtained from ring lattice networks are sinusoids [44], since the eigenvectors of circulant matrices come from the Fourier matrix. Since the resistance distance matrix of a binary ring network is circulant, the signals obtained from \mathbf{R} follow the same rule. Increasing K results in a reduction on the signal's amplitude. Suppose that $K_1 \leq K_2$, then by properties of the resistance distance adding edges between vertices i and j causes $R_{ij}^{(2)} \leq R_{ij}^{(1)}$. Since \mathbf{B} is a Gramian matrix, it follows that $B_{ij}^{(2)} \leq B_{ij}^{(1)}$ [151]. By Weyl's Theorem, $\lambda_m^1 \leq \lambda_m^2$ [146]. Since the signals $X_m(t) = \sqrt{\lambda_m} \cos(\frac{-2\pi mt}{N})$ [44], and the eigenvectors $\cos(\frac{-2\pi mt}{N})$ follow from the Fourier matrix for a ring lattice network, which are independent of the entries of the circulant matrix [152], increasing the node degree K results in a reduction of the signal's amplitude.

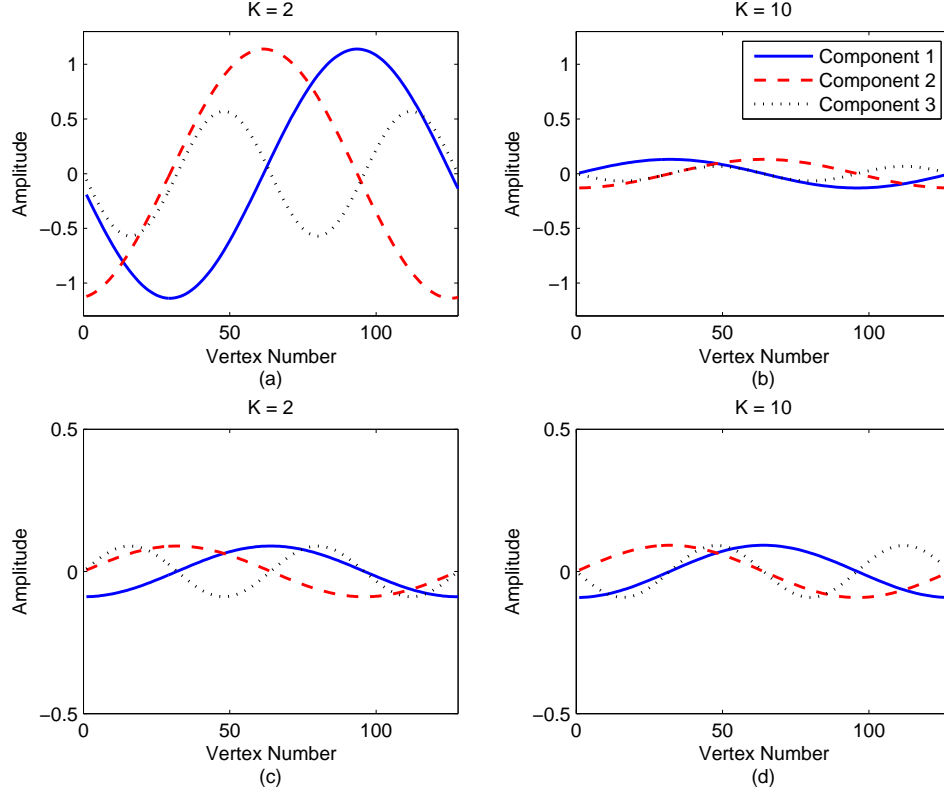


Figure 4.1: Signal representation of a ring lattice network composed of $N = 128$ nodes. Top: Resistance distance; (a) $K = 2$; (b) $K = 10$. Bottom: Distance \mathbf{D} ; (c) $K = 2$; (d) $K = 10$.

Next, we compared both methods in an Erdős-Rényi binary graph for probabilities of attachment p equal to 0.2 and 0.5. For the original distance matrix \mathbf{D} , the signals are random signals, as previously found [46], with amplitudes bounded within the same range for all values of p (Figure 4.2 (c) and Figure 4.2 (d)). On the other hand, signals estimated from \mathbf{R} still exhibit a random structure, with peaks that are inversely proportional to p (Figure 4.2 (a) and Figure 4.2 (b)). The location of these peaks correspond to the nodes with smallest degree, i.e. the largest peak occurs in the first component and corresponds to the node with the smallest degree. In terms of the resistance distance, a node with small degree will have a high resistance distance between it and the remaining nodes in the network. The reduction of the peak's amplitude as the probability of attachment increases follows the previous discussion based on Weyl's Theorem.

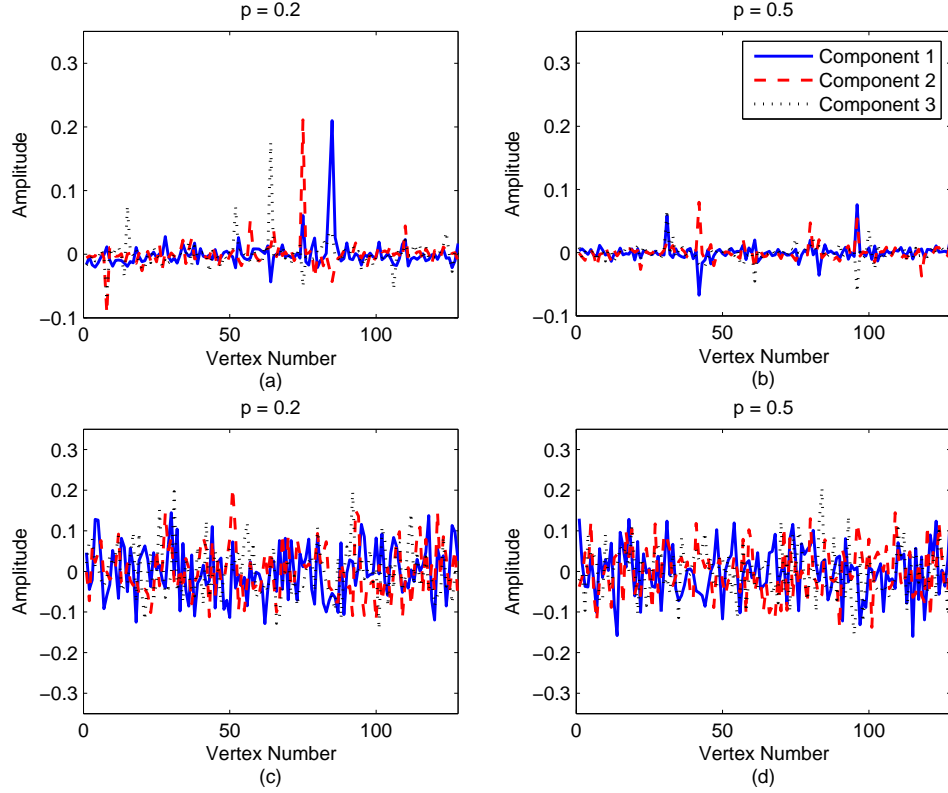


Figure 4.2: Erdős-Rényi network signal representation; (a) and (b), Resistance distance, $p = 0.2$ and $p = 0.5$, respectively; (c) and (d), Distance \mathbf{D} , $p = 0.2$ and $p = 0.5$, respectively.

4.3.5.2 Weighted graphs

The proposed method was also assessed on a weighted stochastic block model consisting of 200 vertices and with fixed probability of attachment, $p = 0.3$. The weights are assigned randomly, uniformly distributed in the interval $[0, 1]$. Figure 4.3 (a) and Figure 4.3 (b) show the signal representations of stochastic block networks with 3 and 4 clusters, respectively. It can be observed from these figures that the first $K - 1$ components reveal the total number of clusters, and the K^{th} component is an impulse. In addition, the size of each cluster can be inferred from the support of the constant regions in the first $K - 1$ components.

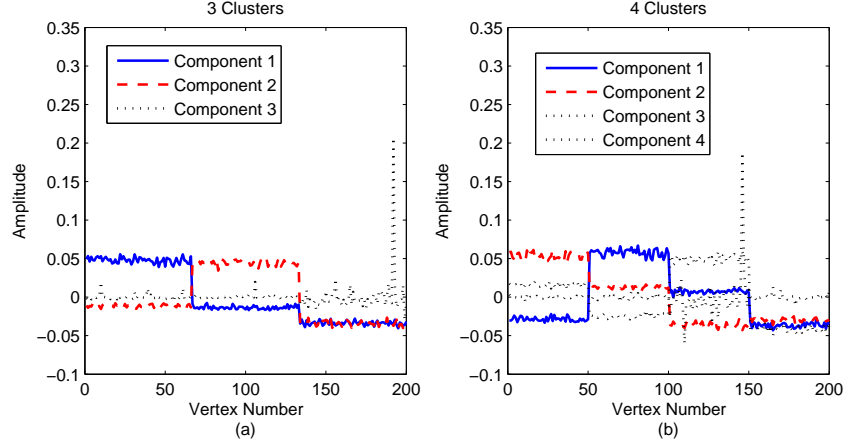


Figure 4.3: Signals constructed from a weighted stochastic block network with probability of attachment $p = 0.3$ using the resistance distance matrix \mathbf{R} . (a) First three components corresponding to a network with 3 blocks; (b) First four components corresponding to a network with 4 blocks.

In addition to the stochastic block network, we investigated the graph to signal transformation of a weighted small-world network. Figure 4.4 shows the signals resulting from a small-world network with average degree $K = 6$, and composed of $N = 128$ vertices. As seen in Figure 4.4 (a), for a network with a low rewiring probability, $p = 0.1$, the resulting signals are sinusoidal signals plus some noise, which increases as p increases (Figure 4.4 (b)). This is consistent with previous works on binary networks [44], where it has been shown that the small-world network is equivalent to a ring network plus noise.

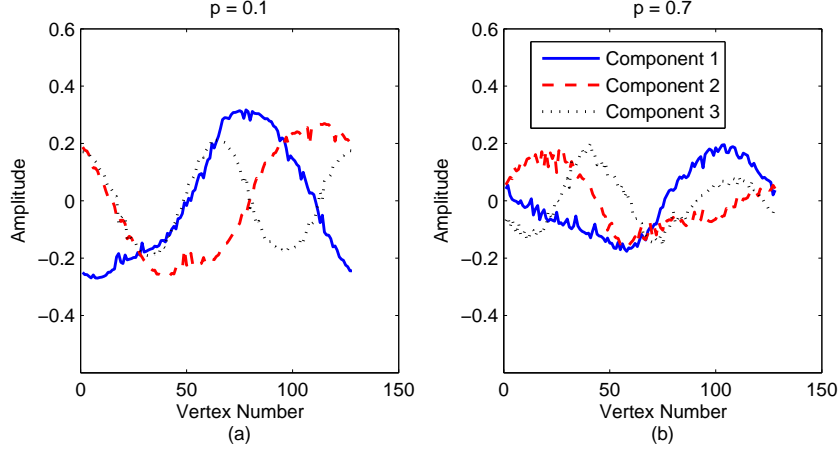


Figure 4.4: Signals constructed from a weighted Small-World network consisting of $N = 128$ nodes and average degree $K = 6$. (a) Rewiring probability $p = 0.1$; (b) Rewiring probability $p = 0.7$.

4.3.5.3 Reconstruction of Weighted Networks

In this section, the reconstruction of networks based on the procedure introduced in Section 4.3.3 is evaluated. The first network considered is a stochastic block model consisting of $N = 150$ nodes, with probability of attachment $p = 0.3$ and 4 clusters. This network is constructed and recovered from its signals \mathbf{X} . The second network is a Erdős-Rényi network consisting of $N = 200$ nodes with probability of attachment $p = 0.5$. Networks were reconstructed for a total of 100 simulations and the reconstruction error is based on the normalized Frobenius norm of the difference between the reconstructed and original adjacency matrices, $\frac{1}{N(N-1)} \|\mathbf{W} - \hat{\mathbf{W}}\|_F$. Table 4.1 shows the reconstruction errors for both networks. This indicates that the proposed reconstruction approach proposed in Section 4.3.3 allows to reconstruct the networks reliably, up to minimal numerical error.

Table 4.1: Reconstruction errors

Network	Error (mean \pm st.dev.)
Stochastic Block	$2.04 \times 10^{-6} \pm 1.43 \times 10^{-6}$
Erdos-Renyi	$2.00 \times 10^{-6} \pm 2.46 \times 10^{-6}$

4.3.5.4 Robustness to network anomalies

In this section, the robustness of the signals obtained from the resistance distance matrix to network anomalies such as missing edges and anomalous edges is assessed. For all the anomalies, the error is computed as the normalized Frobenius norm of the difference between the magnitude spectrum of the original graph and the anomalous graph, $\frac{1}{N(N-1)} \|\mathbf{M} - \hat{\mathbf{M}}\|_F$. By computing the error based on the magnitude spectra of all components instead of the actual signals we avoid misleading error computations that might arise in cases such as ring networks. For ring networks, the corresponding signals are sinusoids and for certain types of anomalies the resulting degraded signal is close to zero, which may result in a small error.

First, we assess the robustness of the method to missing edges. For a weighted ring lattice and a stochastic block network, the weights were uniformly distributed between 0 and 1. Edges are removed at random, while ensuring that the network remains connected. Figure 4.5 (a) and (b) show the error in the signal's spectra for a ring lattice network with $N = 128$ and $N = 256$ nodes, respectively, and the percentage of missing edges ranging from 5% to 20% in increments of 5%. A total of 100 simulations were performed. Figure 4.5 (c) shows the errors for a stochastic block network with 3 clusters as the probability of attachment is varied, which is equivalent to adding more edges to the network, and the percentage of missing edges ranged from 5% to 50% in increments of 5%. As shown in this figure, networks with higher probability of attachment, resulting in more connections, are more robust to the removal of edges.

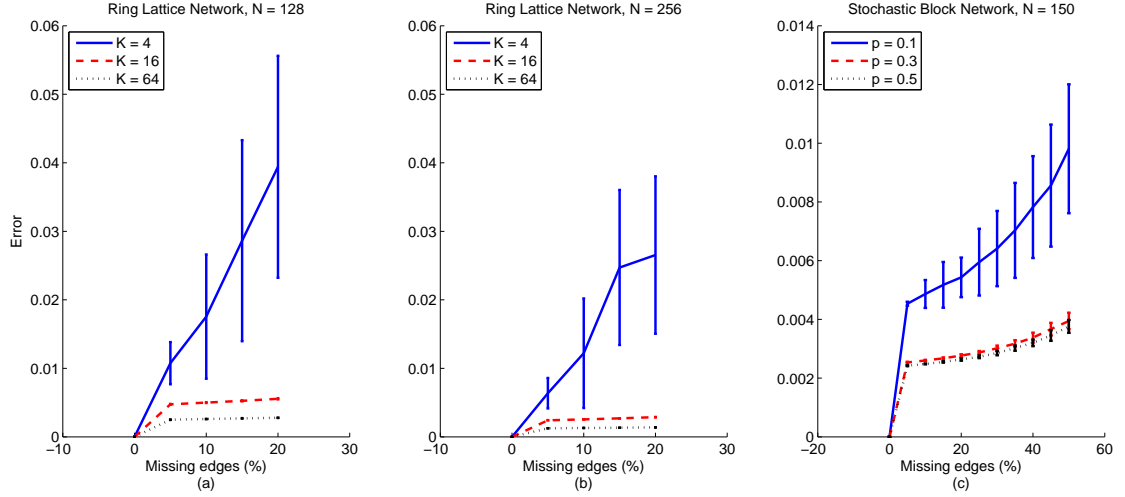


Figure 4.5: Error of the magnitude spectrum. (a) and (b) Ring lattice network with average degree K equal to 4, 16, and 64 consisting of $N = 128$ and $N = 256$ nodes, respectively; (c) Stochastic block network, 3 clusters and probability of attachment $p = 0.1$, $p = 0.3$, $p = 0.5$.

Next, we investigate the robustness of the method when certain edges are anomalous by varying their weight within a certain range. We simulated a weighted ring lattice and a stochastic block network with $N = 128$ nodes, whose edge weights are uniformly distributed within the range $[0.75, 1.25]$. The ring network has an average degree $K = 16$, and the stochastic block network has 3 blocks with a probability of attachment $p = 0.4$. The weight of an anomalous subset of edges is then taken from different amplitude ranges: Range 1: $[0.8, 1.2]$, Range 2: $[0.6, 1.4]$, Range 3: $[0.4, 1.6]$, Range 4: $[0.2, 1.8]$, and Range 5: $[0, 2]$. Figure 4.6 (a) and Figure 4.6 (b) show the error of the magnitude spectrum for the ring lattice network and the stochastic block network, respectively. For both networks, the error increases proportionally to the range of possible values for the anomalous edges and the percentage of anomalous edges.

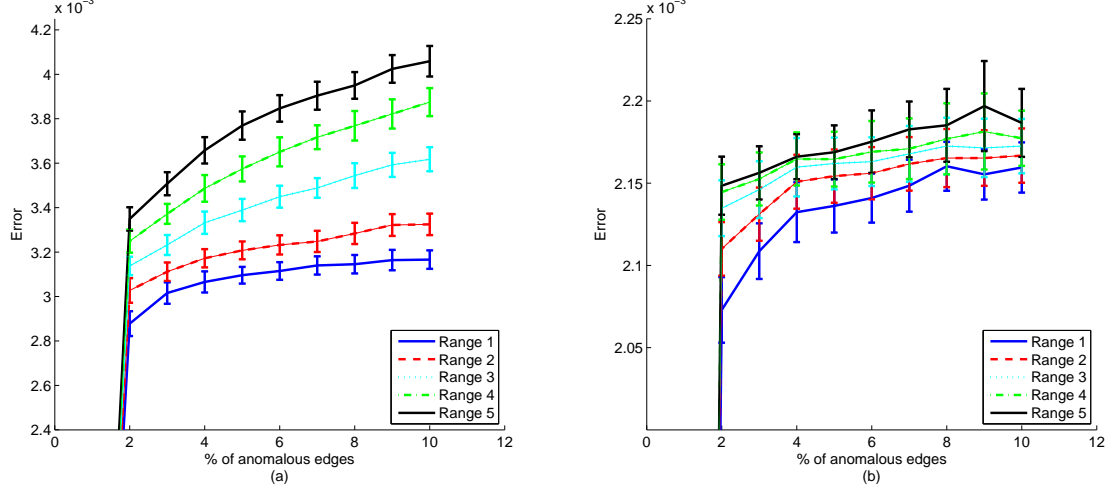


Figure 4.6: Error of the magnitude spectrum from a ring lattice network (a) and a stochastic block network (b) with one anomalous edge whose weight ranged in the intervals $[0.8, 1.2]$, $[0.6, 1.4]$, $[0.4, 1.6]$, $[0.2, 1.8]$, and $[0, 2]$.

4.4 Small-world network characterization

In this section, we propose to characterize the network structure based on the spectrum of the signals \mathbf{X} . We focus on small-world networks and compare our proposed method to the conventional small-world measure for estimating the small-world parameters. Humpires et al. [53] proposed to use the small-world measure for estimating the probability of rewiring, p_r , of real networks with small-world characteristics. They propose to find the p_r that minimizes the error $e(K, p_r, N) = |\sigma(K, p_r, N) - \sigma_{test}|$, where $\sigma(K, p_r, N)$ is the small-world measure of a simulated small-world network with known degree K , probability of rewiring p_r , and N nodes, and σ_{test} is the small-world measure computed from the network under test. However, this estimation method may be affected by various factors related to the small-world measure, as discussed in the introduction.

In this section, we propose to estimate the small-world parameters by correlating the spectral centroid of the signals \mathbf{X} with the spectral centroid of a baseline network. The spectral centroid is the first moment of the normalized power spectral distribution of the signal. Specifically, we denote

the vector of spectral centroids from a baseline simulated network as $\mathbf{c}^B = (c_1^B, c_2^B, \dots, c_N^B)$, where c_i^B refers to the spectral centroid of the i th signal. Similarly, we denote the vector of spectral centroids of the test network as $\mathbf{c}^{test} = (c_1^{test}, c_2^{test}, \dots, c_N^{test})$. Thus, the parameters K and p_r of the small-world network under test are found empirically to maximize the correlation coefficient between the two vectors of centroids, $\rho(K, p_r, N) = \text{corr}(\mathbf{c}^B, \mathbf{c}^{test})$.

Weighted small-world networks with $K = 8$ and varying p_r values were generated 100 times, and then converted into fully connected networks by adding uniformly distributed noise. The weights of the small-world structure were uniformly distributed between $[0, 1]$, while the noise values were uniformly distributed in $[0, 1]$ for Fig. 4.7 (a), and in $[0, 0.25]$ for Fig. 4.7 (b).

The results in Fig. 4.7 show that the proposed method is more accurate than the small-world measure in estimating the probability of rewiring, especially for low p_r within the small-world region. The small-world measure is dependent on the path length and clustering coefficient, which change as more links are added to the network, whereas the spectra of the signals obtained from the graph to signal transformation reflect the underlying small-world structure and is more robust to small changes in the network structure.

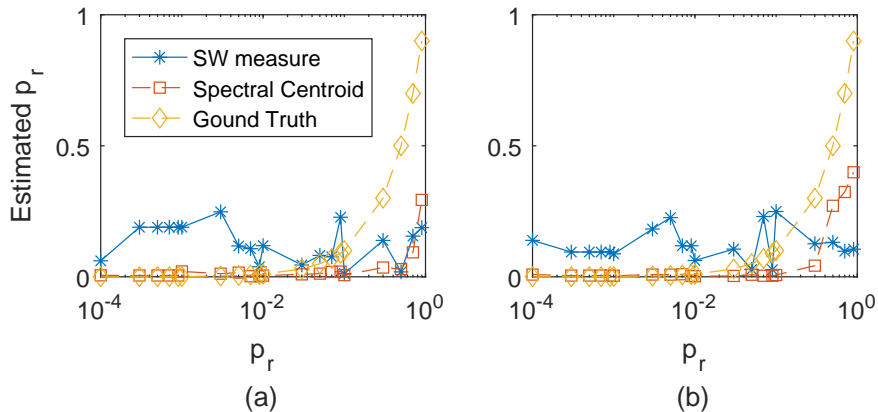


Figure 4.7: Estimated probability of rewiring p_r in weighted small-world networks. Weights of the small world structure are uniformly distributed in the interval $[0, 1]$ and noise values are uniformly distributed in (a) $[0, 1]$, (b) $[0, 0.25]$.

4.5 Graph Entropy based on the graph to signal transform

In this section, we propose a measure for quantifying the structural information of graphs based on the signals obtained from the networks. Since the signals \mathbf{X} convey structural information of the networks, these signals provide alternative distributions to base the graph information theoretic measures on. Specifically, we consider the spectrum and the energy of the signals for computing the Shannon entropy of the networks. The proposed method does not depend on the selection of any parameter nor any graph theoretic measure unlike the previously introduced measures.

The magnitude spectrum of the i th signal is defined as $M_i[f] = |\mathfrak{F}\{x_i\}|^2$, where \mathfrak{F} denotes the discrete Fourier transform, $\mathfrak{F}\{x_i\} = \sum_{n=1}^N x_i[n]e^{-\frac{j2\pi nk}{N}}$. We denote $\mathbf{M} \in \mathbb{R}^{N \times C}$ as the matrix whose columns contain the magnitude spectrum of all signals. The normalized power spectrum of the i th signal is computed as $P_i[f] = \frac{M_i[f]}{\sum_{f=0}^{\lfloor (N-1)/2 \rfloor} M_i[f]}$, where $f = 0, 1, \dots, \lfloor (N-1)/2 \rfloor$ corresponds to discrete frequency bins [153].

We propose to compute the graph entropy based on the normalized power spectrum of $x_i[n]$, $i = 1, 2, \dots, \tilde{C}$, where we consider the \tilde{C} signals with highest energy. This parameter is selected empirically similar to the selection of the total number of factors in Principal Components Analysis (PCA). We propose to use the normalized power spectrum rather than the original signals for entropy computation since computing the Shannon entropy directly on the signals does not necessarily provide information about the network's structural content. For example, for the ring network, the corresponding signals are pure sinusoids 4.3.5 [44], [45], with almost uniform histograms and thus resulting in high entropy. On the other hand, the power spectrum of a sine wave is well localized at a particular frequency (Fig. 4.8 (a)) thus its Shannon entropy is theoretically zero. This is consistent with the intuition that a ring network is deterministic and thus should exhibit low entropy.

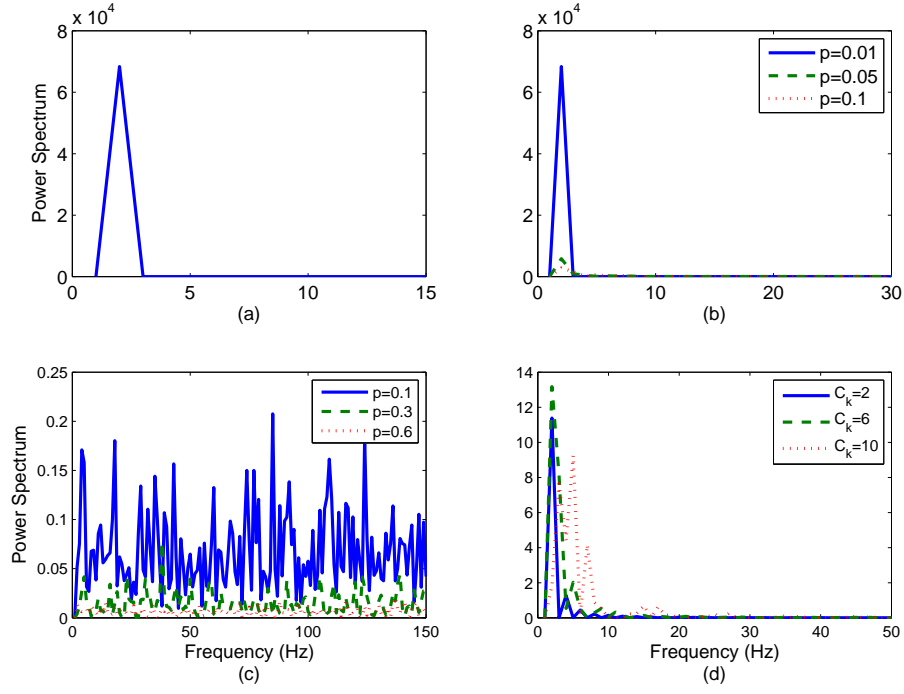


Figure 4.8: Power spectrum of the first graph signal for: (a) Ring network with $K = 4$; (b) Small world network with $p_r = 0.01$, $p_r = 0.05$, and $p_r = 0.1$; (c) Erdős-Rényi network with $p = 0.1$, $p = 0.3$, and $p = 0.6$; (d) Stochastic Block network with $C_k = 2$, $C_k = 6$, and $C_k = 10$ clusters, $N = 300$ nodes for all networks. The frequency axis limits are adjusted in order to better illustrate the spectrum.

The normalized graph entropy for the i th signal is then defined as

$$H_i = -\frac{1}{\log(\lfloor (N-1)/2 \rfloor)} \sum_{f=0}^{\lfloor (N-1)/2 \rfloor} P_i[f] \log(P_i[f]), \quad (4.30)$$

where $i = 1, 2, \dots, \tilde{C}$ [153]. Since (4.30) refers to the Shannon entropy, it is bounded as $0 \leq H_i \leq \log(N^2/2)$. Similar to Shannon entropy, the lower bound is reached when the distribution is an impulse, and the upper bound occurs when the distribution is uniform. In terms of graph structures, the lower bound corresponds to the ring lattice and the upper bound to a random network.

As noted in Fig. 4.8, random networks transform into random signals with a high peak inversely proportional to the probability of attachment p , and hence its spectra are random for all p . In order

to account for the variation in the network entropy as the probability of attachment in random networks varies, we introduce the energy of the signals. We propose to weight each entropy term defined in (4.30) by weights $w_i = \frac{\|\mathbf{x}_i\|_1}{\sqrt{N}\|\mathbf{x}_i\|_2}$, $w_i \in [\frac{1}{\sqrt{N}}, 1]$, using the fact that $\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1 \leq \sqrt{N}\|\mathbf{x}\|_2$. These weights are normalized across signals as $\tilde{w}_i = \frac{w_i}{\sqrt{\tilde{C}}\|\mathbf{w}\|_2}$, where $\mathbf{w} = (w_1, w_2, \dots, w_{\tilde{C}})$. We define the weighted graph spectral entropy (GSE) as

$$GSE = \sum_{i=1}^{\tilde{C}} \tilde{w}_i H_i. \quad (4.31)$$

This definition of network entropy is independent of graph theoretic measures and the eigenspectrum of the adjacency matrix.

Another information theoretic measure of interest in graph theory is graph divergence. In order to assess network dissimilarities, we define the graph Kullback-Leibler divergence based on the graph signal spectrum. First, we define the power spectral density based on the first signal of the graph G_i as $P_{G_i}(f)$. Next, we define the graph Kullback-Leibler divergence between graph G_2 and G_1 as

$$D_{KL}(P_{G_1}||P_{G_2}) = \sum_{f=0}^{\lfloor (N-1)/2 \rfloor} P_{G_1}(f) \log_2 \frac{P_{G_1}(f)}{P_{G_2}(f)}. \quad (4.32)$$

where similar to Kullback-Leibler divergence, $D_{KL}(P_{G_1}||P_{G_2}) \geq 0$. Since the Kullback-Leibler divergence is non-symmetric, we employ the Jensen-Shannon divergence, J -divergence, defined as

$$J(P_{G_1}, P_{G_2}) = \frac{D_{KL}(P_{G_1}||P_{G_2}) + D_{KL}(P_{G_2}||P_{G_1})}{2}. \quad (4.33)$$

We compare the structural entropy obtained from the graph entropies previously defined ($I_{f^{v_1}}$,

$I_{f^{v_2}}$, and \mathcal{H}^1) and the proposed graph-signal based spectral entropy for various networks. Four networks were considered: ring lattice networks with degree $K = 2, 4, 8, 16$, and 32, a small-world network with $K = 8$ and probability of attachment p varying from 0.0001 to 1, a random network with probability of attachment p from 0.1 to 1 in increments of 0.05, and a stochastic block network with $p = 0.1$ and number of blocks $C_k = 3, 5, 7, 9$. All networks have $N = 300$ nodes and the entropies were computed for 200 simulations.

Fig. 4.9 shows the graph entropy for the four networks. As expected, GSE is close to zero for the ring lattice network (Fig.4.9 (a)), whereas the other three entropy measures remain high for all K . In the case of the small-world network (Fig. 4.9 (b)), increasing p_r increases the network structural complexity, and GSE reflects this change in the network structure. Information functionals increase as a function of the probability of rewiring as well, but remain high for low p_r . For small p_r , the network is close to a ring and hence it has low structural complexity. On the other hand, as p_r increases, the network becomes more random and thus its entropy increases. Random networks have high entropy for different probability of attachment p (Fig. 4.9 (c)). Finally, Fig. 4.9 (d) shows the results for a stochastic block network. As reflected by GSE, as the number of clusters increases, the entropy increases.

As observed from the different network structures, the information functionals are not consistent at quantifying the structural complexity in the network. In addition, these require prior knowledge of the network structure in order to select the parameter α . On the other hand, the proposed graph spectral entropy quantifies the network structural information and is sensitive to structural changes in the networks.

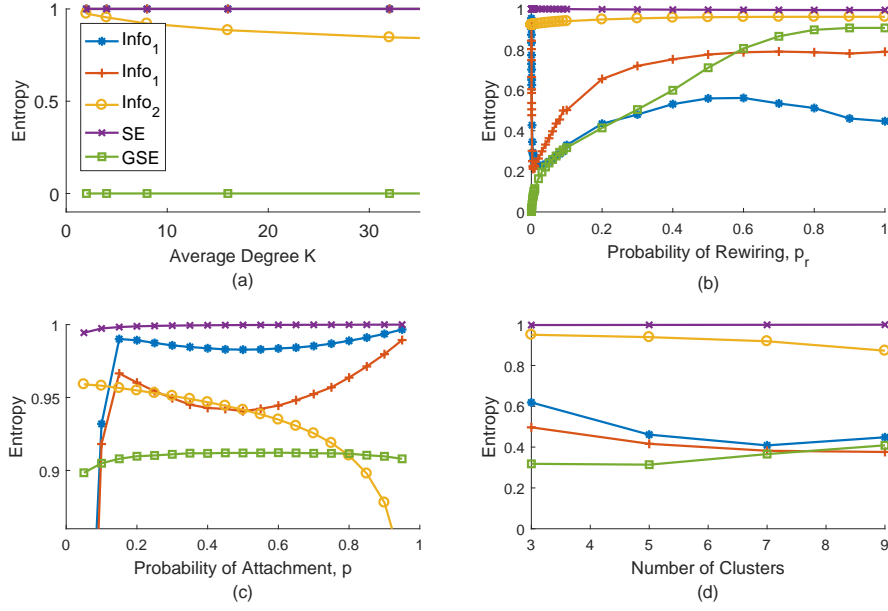


Figure 4.9: Comparison of graph entropy measures for (a) Ring network, $K = 2, 4, 8, 16, 32$ (I_{fV_1} , $\alpha = 0.98$ and $\alpha = 1.03$); (b) Small world network, $K = 4$ and probability of rewiring p_r ranging from 0.0001 to 1 (I_{fV_1} , $\alpha = 0.95$, and $\alpha = 1.05$); (c) Erdős-Rényi network, p from 0.05 to 1 in increments of 0.05 (I_{fV_1} , $\alpha = 0.95$, and $\alpha = 1.1$); (d) Stochastic Block network, $C_k = 3, 5, 7, 9$ (I_{fV_1} , $\alpha = 0.95$, and $\alpha = 1.1$), and $N = 300$ nodes for all networks.

Next, graph divergence between two different binary networks is computed. In the first case, the divergence between two small-world networks, one with mean degree $K = 4$ and $p = 0.0001$ and the other with the same mean degree but different p is assessed. The second case considers the divergence between two stochastic block networks with 3 blocks, the first one with $p = 0.9$ and the second one with varying levels of p . Figure 4.10 (a) and Figure 4.10 (b) show the results for divergence for the small-world and the stochastic block network, respectively. In the small-world network, as p increases, the network becomes less similar to the default network and hence an increase in the divergence is expected. Similarly, in the stochastic block network, as p decreases, the network becomes more random and thus deviates from the default network.

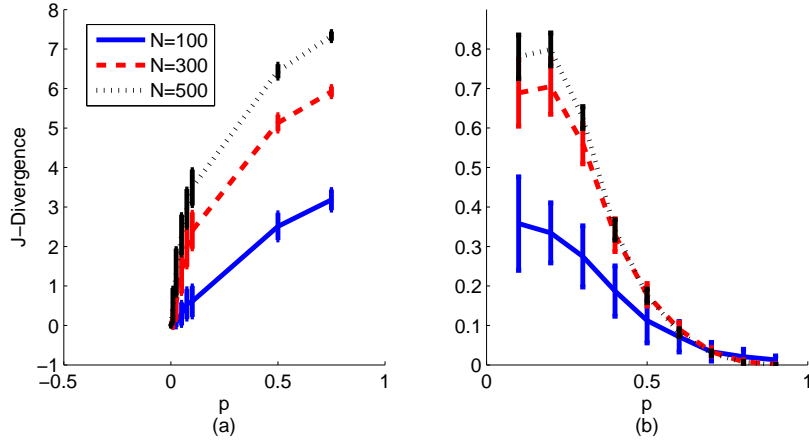


Figure 4.10: Computation of graph divergence between (a) small-world network with $K = 4$ and $p = 0.0001$ and another small-world network with increasing p ; (b) Stochastic Block network with 3 blocks and $p = 0.9$ and another Stochastic Block with 3 blocks and different probability of attachment p .

4.6 Event detection in temporal networks

In this section, we present a method for event detection in temporal networks based on the proposed graph to signal transform. We first introduce background on the tensor decomposition based on the network's adjacency matrix. Next, we describe the proposed method and compare its performance to traditional tensor decompositions based on the adjacency matrix.

4.6.1 Tensor Decompositions for Temporal Networks

Tensor analysis provides a useful tool for revealing the underlying relationships of multilinear data and can be thought of as an extension of PCA and Singular Value Decomposition (SVD) from vector to higher order data. Two major methods for factor analysis of multilinear data are Canonical Polyadic Decomposition or Parallel Factor (CANDECOMP/PARAFAC) and Tucker decomposition [154]. CANDECOMP/PARAFAC is useful in applications where it is desired to factor the data into components that are easily interpretable, such as rank-1 terms, whereas Tucker decompo-

sition is used more often for compression and low-rank projections [139]. In this work, we focus on the PARAFAC decomposition for extracting the temporal profile of dynamic networks since it facilitates its interpretation.

Let $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ be a N th-order tensor. The PARAFAC decomposition approximates the tensor \mathcal{X} as the linear combination of R rank-1 tensors, expressed as [139]

$$\mathcal{X} \approx \sum_{r=1}^R \lambda_r \mathbf{u}_r^{(1)} \circ \mathbf{u}_r^{(2)} \circ \dots \circ \mathbf{u}_r^{(N)}, \quad (4.34)$$

where \circ denotes the outer product. Alternatively, the decomposition in (4.34) can be expressed as

$$\mathcal{X} \approx D \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \dots \times_N \mathbf{U}^{(N)}, \quad (4.35)$$

where $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_R)$, and $\mathbf{U}^{(i)} \in \mathbb{R}^{I_i \times R}$, $i = 1, \dots, N$, are the loading matrices.

In the analysis of temporal networks based on the adjacency matrix, a tensor $\mathcal{X}_a \in \mathbb{R}^{N \times N \times T}$ is constructed, where $\mathcal{X}_a(:, :, t) = \mathbf{A}^{(t)}$ and $\mathcal{X}_a(:, :, t) = \mathbf{W}^{(t)}$, for binary and weighted networks, respectively, and $\mathbf{A}^{(t)} \in \mathbb{R}^{N \times N}$ and $\mathbf{W}^{(t)} \in \mathbb{R}^{N \times N}$ are binary and weighted adjacency matrices at time t , $t = 1, \dots, T$, respectively. The tensor \mathcal{X}_a can be decomposed as

$$\mathcal{X}_a \approx \sum_{r=1}^R \lambda_r \mathbf{u}_r^{(1)} \circ \mathbf{u}_r^{(2)} \circ \mathbf{u}_r^{(3)}, \quad (4.36)$$

where $\mathbf{u}_r^{(1)} \in \mathbb{R}^{N \times 1}$ and $\mathbf{u}_r^{(2)} \in \mathbb{R}^{N \times 1}$ are the same in the case of undirected networks and contain information regarding the node's connectivity, and $\mathbf{u}_r^{(3)} \in \mathbb{R}^{T \times 1}$ is the temporal factor.

4.6.2 Graph to signal transform based event detection

In this section, we propose a tensor decomposition method based on the network's signals for event detection in temporal networks. Let $\mathbf{X}^{(t)} \in \mathbb{R}^{N \times C}$ be the set of signals obtained through CMDS (4.4) at time t . In order to obtain a better insight into the structure of the graph at each time point, we consider the spectra of the graph signals at each time point. We denote the magnitude spectrum of $\mathbf{X}^{(t)}$ as $\mathbf{M}^{(t)}$, and $\mathbf{M}_i^{(t)} \in \mathbb{R}^{N \times 1}, i = 1, \dots, C$, to be the magnitude spectrum of the signal $x_i(n)$ from the network at time t . We then form a tensor $\mathcal{X}_S \in \mathbb{R}^{F \times C \times T}$, where $\mathcal{X}_S(:, :, t) = \mathbf{M}^{(t)}$, and F corresponds to the total number of frequency bins, equal to $\lceil \frac{N}{2} \rceil$ considering only the positive frequencies, C is the total of components obtained from CMDS, and T is the total number of time points.

In order to understand the interactions between the different components across time and frequency, we propose to decompose the tensor \mathcal{X}_S as

$$\mathcal{X}_S \approx \sum_{r=1}^R \lambda_r \mathbf{u}_r^{(1)} \circ \mathbf{u}_r^{(2)} \circ \mathbf{u}_r^{(3)}, \quad (4.37)$$

where $\mathbf{u}_r^{(1)} \in \mathbb{R}^{F \times 1}$ is the spectral factor, $\mathbf{u}_r^{(2)} \in \mathbb{R}^{C \times 1}$ is the components factor, and $\mathbf{u}_r^{(3)} \in \mathbb{R}^{T \times 1}$ is the temporal factor. The tensor decomposition in (4.37) is performed via the MATLAB Tensor Toolbox Version 2.6 [155], [156]. The algorithm enforces nonnegative constraints on the factors and is based on the multiplicative updates of the Nonnegative Matrix Factorization in [157]. The rank of the decomposition, R , is selected according to the core consistency [158].

We compared the proposed method and the tensor decomposition discussed in Section 4.6.1 in the detection of sudden changes in the structure of temporal networks. We generated a temporal weighted small-world network whose edge weights were uniformly distributed between 0.5 and 0.7. The networks consisted of 40 nodes and with $p = 0.01$ for the whole duration of 51 time points,

and created an event by increasing p only at $t = 31$. Four different values of p were considered: $p = 0.05, p = 0.1, p = 0.15$ and $p = 0.2$. A null network with the same properties but without an abrupt event is also constructed for detection analysis, meaning that its size, probability of attachment p and edge weights distribution is the same for all t . For each simulation and network, a total of 200 repetitions were performed, and an Receiver Operating Curve (ROC) for different p values is constructed. In order to construct the ROC, we compared the amplitude of the first component of the temporal mode, $\mathbf{u}_{1_{null}}^{(3)}(t)$ and $\mathbf{u}_{1_{alt}}^{(3)}(t)$, at $t = 31$, corresponding to the null and anomalous network, respectively, to a given threshold. A true detection is defined if the amplitude of $\mathbf{u}_{1_{alt}}^{(3)}(31)$ is greater than the threshold, whereas a false alarm is identified if the amplitude of $\mathbf{u}_{1_{null}}^{(3)}(31)$ is greater than the threshold.

The resulting ROCs for each p are shown in Figure 4.11. It can be observed that for all p , the proposed method (blue) yields a larger area under the curve (AUC) compared to that obtained from the adjacency matrices solely (red) indicating the method's ability to detect sudden changes in the network. Since the magnitude spectrum of the signals does not change considerably as long as the network structure is the same, the temporal mode of the tensor \mathcal{X}_S will reflect only true structural deviations. On the other hand, the entries of the adjacency matrix are distributed within an interval of edge weights and the temporal mode of the tensor \mathcal{X}_a is sensitive to deviations that do not necessarily correspond to structural changes.

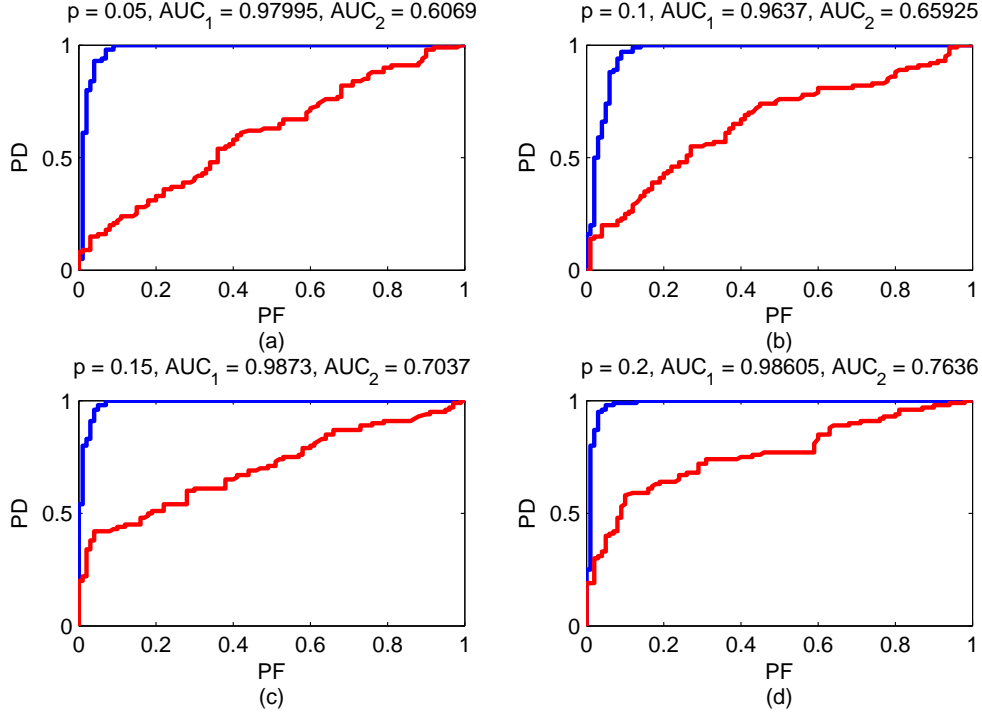


Figure 4.11: Detection of an event consisting of a Small-World network whose probability of attachment p changes from that of the default network ($p = 0.01$) at $t = 31$. ROCs are constructed from the proposed method (blue) and adjacency matrix based method (red) for (a) $p = 0.05$; (b) $p = 0.1$; (c) $p = 0.15$; (d) $p = 0.2$.

4.7 Characterization of Functional Connectivity Networks

In this section, we propose to characterize the network structure based on the spectrum of the signals of functional connectivity networks from the cognitive control experiment described in Chapter 2. Functional connectivity networks are constructed based on the bivariate phase-locking value (PLV) described in Chapter 2 between pairs of electrodes, for both error and correct responses. A network for each subject was constructed by averaging the PLV over the frequency bins corresponding to the theta band, 4-8 Hz, and the ERN interval, 25 – 75 ms. These networks were transformed into signals by using (4.12). The magnitude of the Fourier transform of each component for error, \mathbf{M}^e , and correct, \mathbf{M}^c , responses are shown in Figure 4.12 (a) and Figure 4.12

(b), respectively. As observed in Figure 4.12 (a), the first components of error responses exhibit high energy concentrated in the low frequencies and this energy shifts towards higher frequencies as the component number increases. On the other hand, there is no clear trend in the spectrum corresponding to correct responses. This suggests that functional connectivity networks from error responses have a more organized structure than that of correct responses, which suggests to follow a random, less organized structure.

Next, we assess the similarity of the signal spectra obtained from the error and correct functional connectivity networks to that of a small-world network following the approach proposed in Section 4.4. For both error and correct responses, we computed the spectral centroid for each signal and computed its correlation with that of a small world network for different average degrees K and probabilities of rewiring p_r . Table 4.2 shows the estimated parameters (mean \pm st.dev.) for different time intervals. As observed, FCNs from error responses are characterized with small p_r and K , characteristic of small-world networks, while CRN networks have higher p_r and K , indicative of increased randomness. On the other hand, as observed in Table 4.3, the small-world measure does not reflect such difference. Previous works have reported increased small-worldness for ERN compared to CRN [116]. Therefore, the proposed approach can serve as an alternative method for the characterization of FCN structure in distinct cognitive states, and furthermore, estimate network parameters.

Table 4.2: Estimated small-world parameters.

Interval	Estimated small-world parameters			
	p_r		k	
	ERN	CRN	ERN	CRN
-25 - 0	0.0113 ± 0.0152	0.2897 ± 0.3602	2.1111 ± 0.4714	9.2222 ± 6.9668
0 - 25	0.0060 ± 0.0044	0.3707 ± 0.3910	2 ± 0	13 ± 6.5530
25-50	0.0119 ± 0.0213	0.3474 ± 0.3636	2 ± 0	9.7778 ± 6.8561
50-75	0.0085 ± 0.0137	0.3533 ± 0.3801	2 ± 0	10.2222 ± 7.9377
75-100	0.0065 ± 0.0065	0.3622 ± 0.3897	2 ± 0	11.7778 ± 6.9583
100-125	0.007 ± 0.0111	0.2734 ± 0.3740	2 ± 0	12.5556 ± 6.2046

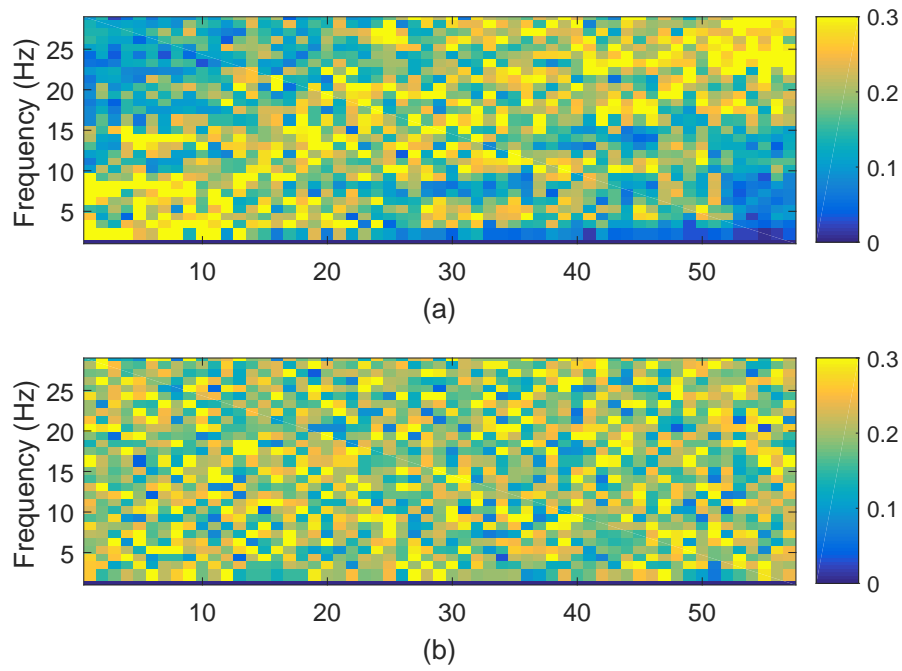


Figure 4.12: Magnitude Spectrum for each signal obtained through network to signal transformation for (a) Error responses; (b) Correct responses.

Table 4.3: Small-world measure.

Interval	σ	
	ERN	CRN
-25 - 0	1.165 ± 0.2827	1.1677 ± 0.2833
0 - 25	1.1635 ± 0.2823	1.1662 ± 0.2829
25 - 50	1.1689 ± 0.2823	1.1717 ± 0.2844
50-75	1.1653 ± 0.2826	1.1681 ± 0.2833
75-100	1.166 ± 0.2828	1.1687 ± 0.2835
100-125	1.1671 ± 0.2829	1.1698 ± 0.2836

4.7.1 Assessment of Graph Information Theoretic Measures in Functional Connectivity Networks

Table 4.4 shows the entropy results (mean \pm st.dev.) for the ERN and CRN FCNs over six different time intervals. For all intervals, FCNs from correct responses show higher entropy than FCNs from error responses and this difference is significant for all intervals ($p < 0.05$, Wilcoxon rank-sum test). This is consistent with the fact that the error-related negativity is associated with increased synchronization which results in less random networks and hence lower network entropy. In addition, the slight increase in the network entropy within the ERN interval (25-75 ms) can be related to results from a previous study [159], where it has been shown that ERN is associated with increased segregation within the FCN resulting in more clusters, i.e. higher entropy in the network organization.

In addition, we correlate the entropy results of each subject with behavioral measures relevant to the cognitive control experiment. In particular, we considered the post-error slowing (PES) and post-error accuracy (PEA). PES is computed as correct response time after error responses minus correct response time after correct responses, and PEA is computed as the accuracy after error responses minus the accuracy after correct responses. Table 4.5 shows the correlation between

Table 4.4: Graph entropy from cognitive control FCNs.

	Entropy (mean \pm st.dev.)		
Interval (ms)	ERN	CRN	p-value
-25-0	0.8511 ± 0.0197	0.8738 ± 0.0072	0.0001
0-25	0.8575 ± 0.0138	0.8726 ± 0.0073	0.0042
25-50	0.8594 ± 0.0130	0.8724 ± 0.0078	0.0048
50-75	0.8542 ± 0.0157	0.8713 ± 0.0095	0.0090
75-100	0.8510 ± 0.0168	0.8738 ± 0.0090	0.0003
100-125	0.8513 ± 0.0169	0.8723 ± 0.0066	0.0002

Table 4.5: Correlations between graph entropy and behavioral measures from cognitive control.

	PES		PEA	
Interval (ms)	ERN	CRN	ERN	CRN
-25-0	0.1867	0.0325	-0.1086	0.0041
0-25	-0.0879	-0.0626	-0.269	0.054
25-50	0.3358	-0.2744	-0.0406	0.1959
50-75	0.2772	-0.0518	-0.1373	0.0299
75-100	0.3807	0.4019	-0.1183	0.1128
100-125	0.2976	-0.2697	0.2504	-0.128

the FCN entropy and behavioral measures over six different time intervals. As observed, PES is positively correlated with the graph entropy of error FCNs over the ERN interval, while the FCN entropy of correct responses is negatively correlated. This result follows previous studies showing an inverse relationship between PES and increased synchrony in the error-related activity [160], [161], and hence, PES is directly proportional to the network entropy during the ERN.

4.8 Conclusions

In this chapter, a new network to signal transformation based on the resistance distance has been proposed for both binary and weighted networks. This is the first deterministic graph to signal transform proposed for both binary and weighted networks. This transform is also shown to guarantee the reconstruction of the networks from their signals. Transforming graphs into signals

provides the benefit of applying traditional signal processing measures to these signals in order to assess certain properties of the networks. Along those lines, the proposed graph to signal transform served as the basis for the introduction of graph information theoretic measures, an approach for small-world network characterization, and an event detection approach in temporal networks proposed in this chapter.

First, we showed theoretical properties of the proposed method and the resistance distance matrix. The graph to signal transformation of various well-known network structures was illustrated for both binary and weighted networks. For binary networks, the proposed method reveals structural attributes of the graphs not perceivable by a previously proposed distance matrix. Furthermore, analysis of perturbations in the network and how these are reflected in the proposed graph to signal transform were presented as well. In addition, through simulations, we demonstrated the behavior of our proposed technique to network anomalies.

Second, an approach for the computation of the structural information content of graphs using the network to signal transform was introduced. The proposed method considers the normalized power spectrum of the graph signal with the highest energy as a probability distribution to be employed in the computation of Shannon entropy and Kullback-Leibler divergence. This method is advantageous over current graph information theoretic measures due to its independence from arbitrary parameters. The results from simulated networks illustrate that the proposed method yields a reliable characterization of the structural information of the graph. Furthermore, it allows for the quantification of the structural information of functional connectivity networks and reflects differences between two cognitive states during the particular time interval of interest.

In addition, we introduced a graph-signal transformation based approach for detecting events in temporal networks. The method is based on factor analysis of a tensor formed from the spectra of the different signal components across time, where the factors along the time mode reveal the

change to the network structure. We compared the temporal factors from the tensor decomposition of the proposed approach to the factors from the tensor decomposition of the adjacency matrices over time. By comparing the first component of the temporal mode at the time where the event occurred, we showed from ROC analysis that the proposed method results in a higher detection rate of abrupt structural changes in temporal networks when compared to the tensor decomposition of adjacency matrices over time.

Lastly, it was shown how the proposed method can characterize the structural properties of functional connectivity networks under different cognitive states. Following a priori knowledge suggesting that functional connectivity networks behave as small-world networks, it was shown how the spectral centroid of the functional connectivity networks signals from the proposed graph to signal transform correlate to the spectral centroid of the signals from small-world networks, for different network parameters. From these results, it was shown that functional connectivity networks from error responses are highly correlated to a small-world network, whereas the networks for the correct responses are less correlated.

Chapter 5

Dynamic Graph Fourier Transform

Recent research in signal processing over graphs has provided the tools for processing signals defined on irregular domains such as graphs [49]. In many applications, such as social networks, sensor networks, energy networks, and brain networks, among others, signals lie on the set of vertices of the network. Other fields where data is defined on irregular domains include data defined on manifolds and irregularly shaped domains [162], such as cells in histological images [163], and data based on point clouds. Recently, it has been shown [164] how signal processing methods adapted to signals on graphs such as filtering and Fourier transform defined in the context of signal processing over graphs provide insights about learning processes in the brain.

Various transforms from signal processing have been adapted to the graph domain to analyze the spectral content of signals over graphs. The first is the graph Fourier transform (GFT), which aims to compute the Fourier transform of a signal defined on the vertices of a graph by employing a basis obtained from the network's adjacency matrix [165] or Laplacian matrix [49]. Another transform defined on graph signals is the windowed graph Fourier transform [166], which considers the nonstationarity of the graph signals and transforms them to the vertex-frequency domain. In order to define a windowed graph Fourier transform over graph signals, in [166] the authors define generalized convolution, translation, and modulation operators for signals on graphs. In addition, a wavelet transform for graph signals has been developed in [162], known as spectral graph wavelets

since it is based on spectral graph theory. By doing this, scaling is defined from the eigenfunctions of the graph Laplacian and avoid its computation over irregular domains. Recently, the joint time-vertex Fourier transform [167] was proposed for graph signals evolving over time. The joint time-vertex Fourier transform is found by first computing the GFT along the graph dimension and the discrete Fourier transform along the time-domain. In addition, the dynamic graph wavelet transform [168] has been proposed for the case when the time-vertex domain is dynamic. However, in all of these approaches the underlying graphs are stationary.

In some applications, such as functional connectivity networks in the brain, the underlying network structure varies over time [29], [169]. This requires the adaptation of the previously mentioned graph signal transforms in order to consider the nonstationary network structure. For example, in the case of the graph Fourier transform, the adjacency matrix or the Laplacian matrix of the network changes for each time instance, and a unique spectral representation is not possible. Therefore, there is no unique definition of frequency across time as the graph evolves. One alternative would be averaging of the adjacency matrix or the Laplacian [51]. However, averaging does not necessarily find the optimal subspace across time. This problem has been previously addressed by defining a common Laplacian across time, where a common subspace was found by means of Grassmann manifolds [50]. There, the authors used this common subspace in the definition of a dynamic graph Fourier transform. However, the accuracy of a common subspace is compromised as the number of time points increases. In addition, when the network structure is nonstationary, the size of the window in which the common subspace is found should be determined by the characteristics of the network.

In this chapter, we propose a dynamic graph Fourier transform (dGFT) for which a common subspace estimate is found by means of tensor decomposition. The temporal network adjacency matrices or Laplacian matrices over time constitute a 3-way tensor. The Tucker decomposition of

this tensor results in orthonormal component matrices which define the basis of the time-varying Laplacian operator. The obtained basis and the corresponding subspace are optimal in the sense of finding the best low-rank approximation to the Laplacians across time.

This chapter is organized as follows. Section 5.1 introduces background on graph signal processing and tensor decompositions. Section 5.2 presents the proposed tensor based dGFT. Section 5.3 presents results on simulated graph signals and dynamic functional connectivity networks (dFCNs) constructed from cognitive control EEG experiment. Section 5.4 presents the conclusions and future work.

5.1 Background

5.1.1 Graph Signal Processing

A graph signal $f : V \rightarrow \mathbb{R}$ is defined on the vertices of the graph G . It is represented by a vector $\mathbf{f} \in \mathbb{R}^{N \times 1}$, and the i^{th} element of this vector corresponds to the signal at vertex v_i [165]. Thus, signal amplitudes at each node define a graph signal, and it is indexed by the graph nodes. Since the signals are defined over the graph nodes, the underlying network structure plays an important role in the definition of transformations of the signals \mathbf{f} . In particular, the adjacency matrix or the graph Laplacian are employed in the analysis of graph signals.

In this work, we focus on the graph Laplacian. The Laplacian \mathbf{L} is a positive semidefinite and real matrix and thus has a complete set of orthonormal eigenvectors $\{\mathbf{u}_l\}_{l=0,1,\dots,N-1}$, and eigenvalues $\{\lambda_l\}_{l=0,1,\dots,N-1}$, $0 = \lambda_0 \leq \lambda_1 \leq \dots \leq \lambda_{N-1}$. Therefore, it admits the eigendecomposition

$$\mathbf{L} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T, \quad (5.1)$$

where $\mathbf{U} = [\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_{N-1}]$, and $\Lambda = \text{diag} \{\lambda_0, \lambda_1, \dots, \lambda_{N-1}\}$.

The spectrum of the graph Laplacian has been widely used in applications such as clustering and spectral matching. Recently, the Laplacian eigenvectors $\{\mathbf{u}_l\}_{l=0,1,\dots,N-1}$ have been proposed as the Fourier basis for the graph Fourier transform [49]. As in the classical Fourier analysis, the eigenvectors of \mathbf{L} provide a notion of frequency since, for connected graphs, the eigenvector corresponding to the smallest eigenvalue is constant, equal to $\frac{1}{\sqrt{N}}$. As the frequency λ_i increases, the eigenvectors oscillate more rapidly.

The graph Fourier transform (GFT) of \mathbf{f} defined on the graph vertices V is given by [49]

$$\hat{f}(\lambda_l) = \langle \mathbf{f}, \mathbf{u}_l \rangle = \sum_{i=1}^N f(i) \mathbf{u}_l^*(i), \quad (5.2)$$

where $\mathbf{u}_l, l = 0, 1, \dots, N-1$ correspond to the eigenvectors of the graph Laplacian, and the frequencies are indexed by its corresponding eigenvalues. The inverse graph Fourier transform (iGFT) is obtained by

$$f(i) = \sum_{l=0}^{N-1} \hat{f}(\lambda_l) \mathbf{u}_l(i). \quad (5.3)$$

In matrix form, the graph Fourier transform and its inverse are given as

$$\hat{\mathbf{f}} = \mathbf{U}^T \mathbf{f}, \quad (5.4)$$

and

$$\mathbf{f} = \mathbf{U} \hat{\mathbf{f}}, \quad (5.5)$$

respectively.

The graph Laplacian is the discrete difference operator, and for every vector $\mathbf{y} \in \mathbb{R}^N$ it satisfies

$$\mathbf{y}^T \mathbf{L} \mathbf{y} = \frac{1}{2} \sum_{i,j=1}^N W_{ij} (y_i - y_j)^2, \quad (5.6)$$

where W_{ij} is the (i, j) entry of the graph adjacency matrix. In the context of graph signal processing, the graph Laplacian quadratic form, $S_2(\mathbf{f}) = \mathbf{f}^T \mathbf{L} \mathbf{f}$ is referred to as the smoothness of the signal \mathbf{f} with respect to the Laplacian \mathbf{L} . The GFT can be interpreted in terms of the smoothness of the graph signals the GFT of a smooth signal will occupy the low frequencies λ in the spectrum, and will have a small $S_2(\mathbf{f})$. This occurs when two neighbor vertices are connected by a edge with large weight and the signal \mathbf{f} at those vertices has similar values.

In order to filter the graph signals, low-pass $\hat{\mathbf{h}}_{Lk}$, band-pass $\hat{\mathbf{h}}_{Bk}$ and high-pass $\hat{\mathbf{h}}_{Hk}$ graph filters are defined as

$$\hat{\mathbf{h}}_{Lk} = \mathbb{I}_{\{k < K_L\}}, \quad (5.7)$$

$$\hat{\mathbf{h}}_{Bk} = \mathbb{I}_{\{K_L \leq k < K_L + K_B\}}, \quad (5.8)$$

$$\hat{\mathbf{h}}_{Hk} = \mathbb{I}_{\{K_L + K_B \leq k\}}, \quad (5.9)$$

where K_L and K_B correspond to the cut-off frequencies and \mathbb{I} is the indicator function. The graph signals \mathbf{f} are filtered as $\hat{\mathbf{f}}_L = \hat{\mathbf{H}}_L \hat{\mathbf{f}}$, $\hat{\mathbf{f}}_B = \hat{\mathbf{H}}_B \hat{\mathbf{f}}$, and $\hat{\mathbf{f}}_H = \hat{\mathbf{H}}_H \hat{\mathbf{f}}$, where $\hat{\mathbf{H}}_L = \text{diag}(\hat{\mathbf{h}}_{Lk})$, $\hat{\mathbf{H}}_B = \text{diag}(\hat{\mathbf{h}}_{Bk})$, and $\hat{\mathbf{H}}_H = \text{diag}(\hat{\mathbf{h}}_{Hk})$ are the filters. Taking the iGFT we obtain the filtered graph signals \mathbf{f}_L , \mathbf{f}_B , and \mathbf{f}_H .

5.1.2 Tucker Decomposition

Let $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ be a 3rd-order tensor. The tensor \mathcal{X} can be decomposed by means of Tucker decomposition as

$$\mathcal{X} \approx \mathcal{C} \times_1 \mathbf{B}^{(1)} \times_2 \mathbf{B}^{(2)} \times_3 \mathbf{B}^{(3)}, \quad (5.10)$$

where $\mathcal{C} \in \mathbb{R}^{I \times J \times K}$ is the core tensor and the factor matrices $\mathbf{B}^{(1)} \in \mathbb{R}^{I \times I}$, $\mathbf{B}^{(2)} \in \mathbb{R}^{J \times J}$, and $\mathbf{B}^{(3)} \in \mathbb{R}^{K \times K}$ are orthogonal. The matrices $\mathbf{B}^{(1)}$, $\mathbf{B}^{(2)}$, and $\mathbf{B}^{(3)}$ can be obtained as the left singular vectors of $\mathbf{X}_{(1)} \in \mathbb{R}^{I \times JK}$, $\mathbf{X}_{(2)} \in \mathbb{R}^{J \times KI}$, and $\mathbf{X}_{(3)} \in \mathbb{R}^{K \times IJ}$, respectively [170], illustrated in Figure 5.1. For large matrices, the factor matrices $\mathbf{B}^{(i)}$ can be obtained as the eigenvectors of $\mathbf{X}_{(i)} \mathbf{X}_{(i)}^T$. It is important to emphasize that the Tucker decomposition is not unique. The Tucker decomposition based on matrix unfolding discussed before is referred to as Tucker 1 [171]. Other decompositions are referred to as Tucker 2 and Tucker 3, and these allow rank reduction in more than one mode, which should be specified from the user. Some observations about Tucker 3 [171] include its flexibility since the core allows interactions between factors in different modes, and that it cannot determine the component matrices uniquely.

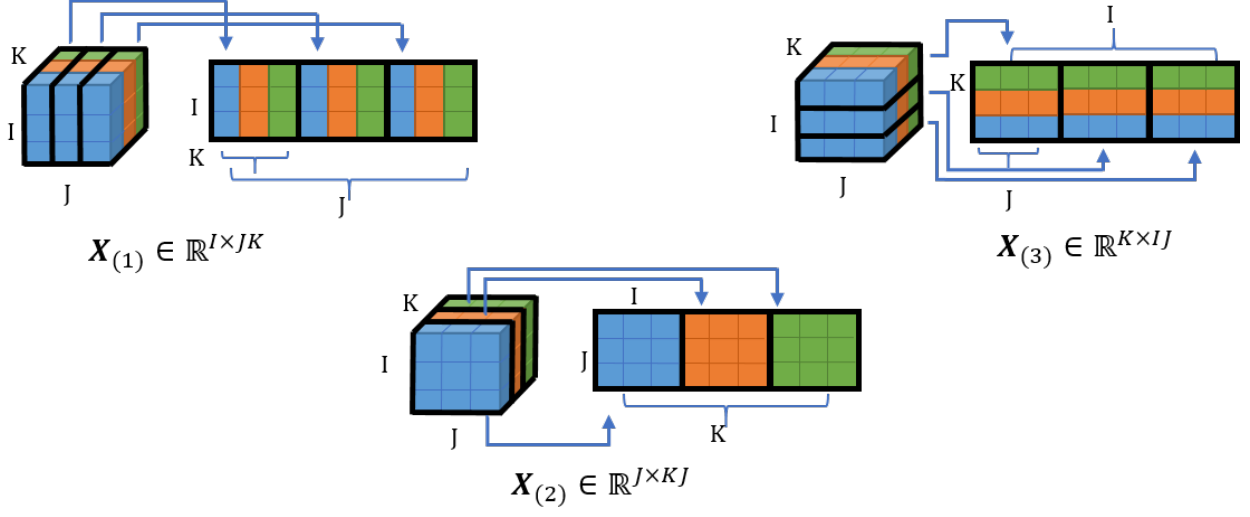


Figure 5.1: Tensor unfolding.

5.2 Dynamic Graph Fourier Transform on Temporal Networks

Consider the dynamic network $G^{(t)} = (V, E^{(t)}, A^{(t)})$, $t = 1, 2, \dots, T$, to be a time-varying network whose edges vary with time and the vertex set remains constant. The adjacency matrices $\mathbf{A}^{(t)}$ over time constitute the 3-way tensor $\mathcal{A} \in \mathbb{R}^{N \times N \times T}$, where N is the total number of vertices, T is the total number of time points, and $\mathcal{A}(:, :, t) = \mathbf{A}^{(t)}$. Similarly, we can define the 3-way tensor $\mathcal{D} \in \mathbb{R}^{N \times N \times T}$ from the degree matrices $\mathbf{D}^{(t)}$ over time, where $\mathcal{D}(:, :, t) = \mathbf{D}^{(t)}$. Since in traditional GFT the eigenvectors of the Laplacian define the basis for the transform, we use the same idea to find the common subspace of the Laplacians, $\mathbf{L}^{(t)}$, across time. Some possible approaches to combining multiple Laplacians include averaging, weighted averaging [172, 173] and a more recent optimization framework based on a maximum likelihood criterion [174]. In order to find the common subspace, we define the 3-way tensor from the Laplacians of the time varying graph as $\mathcal{L} \in \mathbb{R}^{N \times N \times T}$, where $\mathcal{L}(:, :, t) = \mathcal{D}(:, :, t) - \mathcal{A}(:, :, t)$, and find the subspace information through Tucker decomposition as

$$\mathcal{L} \approx \mathcal{C} \times_1 \mathbf{U} \times_2 \mathbf{U} \times_3 \mathbf{V}, \quad (5.11)$$

where $\mathcal{C} \in \mathbb{R}^{N \times N \times T}$ is the core tensor and $\mathbf{U} \in \mathbb{R}^{N \times N}$, and $\mathbf{V} \in \mathbb{R}^{T \times T}$ are the orthogonal factor matrices along the connectivity and time modes. Due to the symmetry of \mathcal{L} along the first and second modes, the corresponding factor matrices are identical.

We propose to consider the left singular vectors of the matrix $\mathbf{L}_{(1)} \in \mathbb{R}^{N \times NT}$, $\mathbf{u}_l, l = 0, 1, \dots, N-1$, as the common basis to be employed in the graph Fourier transform of the time-varying network $G^{(t)}$. This procedure avoids the need for finding a common Laplacian matrix as an intermediate step and uses the orthogonal basis that spans the connectivity mode across all time. A common Laplacian $\mathbf{L}_T = \mathbf{U}\Sigma\mathbf{U}^T$ is computed, where $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_N)$ are the singular values of $\mathbf{L}_{(1)}$.

Let $\mathbf{f}^{(t)} \in \mathbb{R}^{N \times 1}$ be the signal defined on the vertices V at time t . The dGFT of $\mathbf{f}^{(t)}$ is then given by

$$\hat{f}^{(t)}(\lambda_l) = \langle \mathbf{f}^{(t)}, \mathbf{u}_l \rangle = \sum_{i=1}^N f^{(t)}(i) \mathbf{u}_l^*(i), \quad (5.12)$$

where \mathbf{u}_l is the l th column of \mathbf{U} in (5.11). Note that this transform is a function of time and vertex frequency. We also define the dGFT based on the eigenvectors of the average Laplacian matrix \mathbf{L}_A over time, $\hat{\mathbf{L}}_A = \frac{1}{T} \sum_{t=1}^T \mathbf{L}^{(t)}$ and denote it as dGFT_L.

5.3 Results

In this section, we first compare the dGFT obtained from the proposed method and the average Laplacian. Second, we assess the dynamic graph Fourier transform on dynamic functional connec-

tivity networks for the cognitive control study.

5.3.1 Simulations

We simulated a weighted ring lattice network with $N = 100$ nodes, with average degree $K = 4$ for $T = 80$. At each time instance, the edge weights were selected randomly from the interval $[0.75, 1]$ in order to simulate slight variations present in real networks. The signal $\mathbf{f}^{(t)} \in \mathbb{R}^{N \times 1}$ is defined as

$$\mathbf{f}^{(t)} = \begin{cases} \mathbf{v}_{10}^{(5)}, & t = 0, \dots, 20, \\ \mathbf{v}_{10}^{(5)} + \mathbf{v}_{40}^{(12)} + \mathbf{v}_{60}^{(50)}, & t = 21, \dots, 50, \\ \mathbf{v}_5^{(75)} + \mathbf{v}_{40}^{(12)} + \mathbf{v}_{80}^{(35)}, & t = 51, \dots, T, \end{cases} \quad (5.13)$$

where $\mathbf{v}_i^{(t)}$ is the i th eigenvector of the network Laplacian at time t .

Figure 5.2 shows the results from the proposed dGFT (5.12). The results from dGFT_A in this simulation are similar and not shown. In order to facilitate the interpretation of the results, the frequency axis is normalized by the largest eigenvalue. As expected, in the interval $0 \leq t \leq 20$ the frequency content of the network corresponds to $\mathbf{v}_{10}^{(5)}$, which extends until $t = 50$. In the second interval, $21 \leq t \leq 50$, there are in addition the frequency components corresponding to eigenvectors 40 and 60. Finally, during the last interval there are components corresponding to eigenvectors 5, 40, and 80.

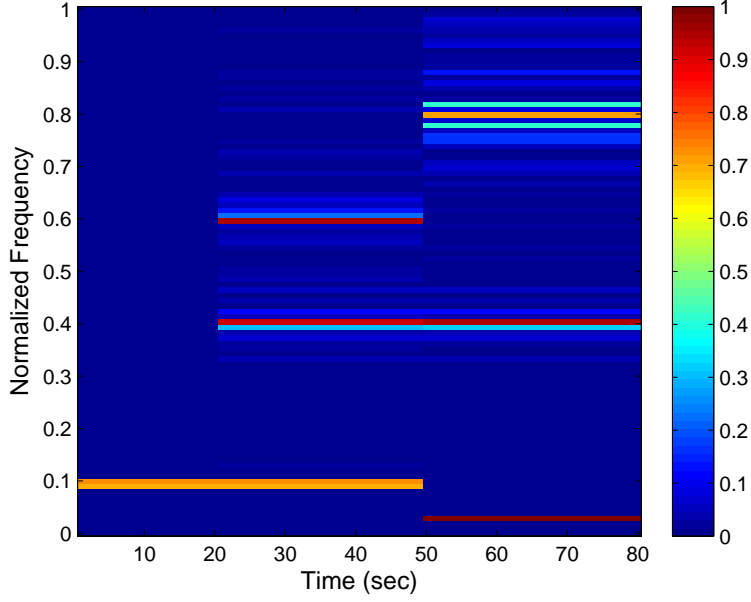


Figure 5.2: DGFT of a ring network with $N = 100$ nodes and $K = 4$ over $T = 80$ seconds. The graph signal is composed of different components over time, which are extracted by the proposed method.

Next, we assess how the proposed common subspace is affected by changes in the network structure. A dynamic Erdős-Rényi graph is generated for $T = 8$ time points, with $N = 60$ nodes and probability of attachment $p = 0.75$. The weights are varied randomly in the interval $[0.75, 1]$ (Interval 1) for the first T_1 time points and in $[0.25, 0.5]$ (Interval 2) for the remaining $T - T_1$ time points. We consider three different T_1 : $T_1 = 2$, $T_1 = 4$, $T_1 = 6$. This simulation is performed for two different graph signals $\mathbf{f}_1^{(t)}$ and $\mathbf{f}_2^{(t)}$, $t = 1, \dots, T$, with their amplitudes uniformly distributed in $[0.9, 1]$ and $[-1, 1]$, respectively. In order to assess the performance of the proposed method, we computed the average error within a time interval as $\frac{1}{T'} \sum_{t=t_1}^{T'} \|\tilde{\mathbf{L}} - \mathbf{L}^{(t)}\|_F$, where T' is the duration of the interval and $\tilde{\mathbf{L}}$ is the estimated common Laplacian over all time, either \mathbf{L}_T or \mathbf{L}_A . We also computed the coding gain, $CG(\mathbf{U}, \mathbf{L})$ which quantifies how the common subspace diagonalize the Laplacian pseudoinverse at time t , $\mathbf{L}^{\dagger(t)}$, and is computed as

Table 5.1: Performance of the proposed \mathbf{L}_T and \mathbf{L}_A .

	Error				CG			
	Interval 1		Interval 2		Interval 1		Interval 2	
T_1	\mathbf{L}_T	\mathbf{L}_A	\mathbf{L}_T	\mathbf{L}_A	\mathbf{L}_T	\mathbf{L}_A	\mathbf{L}_T	\mathbf{L}_A
2	116.367	132.798	79.985	74.662	0.924	0.924	0.935	0.935
4	74.220	91.021	89.761	89.306	0.923	0.923	0.935	0.935
6	43.004	52.428	92.873	91.602	0.923	0.923	0.935	0.935

Table 5.2: MSE of the smoothness of $\mathbf{f}_1^{(t)}$ from \mathbf{L}_T and \mathbf{L}_A .

	MSE			
	Interval 1		Interval 2	
T_1	\mathbf{L}_T	\mathbf{L}_A	\mathbf{L}_T	\mathbf{L}_A
2	1.039 ± 0.132	1.190 ± 0.152	0.949 ± 0.123	0.688 ± 0.088
4	0.853 ± 0.093	1.071 ± 0.117	1.288 ± 0.142	1.069 ± 0.117
6	0.502 ± 0.058	0.687 ± 0.079	1.288 ± 0.150	1.180 ± 0.137

$$CG(\mathbf{U}, \mathbf{L}^{\dagger(t)}) = \left(\frac{\prod_{i=1}^N \mathbf{L}_{ii}^{\dagger(t)}}{\prod_{i=2}^N (\mathbf{U}^t \mathbf{L}_{ii}^{\dagger(t)} \mathbf{U})_{ii}} \right)^{1/N}. \quad (5.14)$$

In addition, we computed the mean squared error (MSE) of the smoothness from each of the common Laplacians, $MSE(S_{2\mathbf{L}_T}(\mathbf{f}^{(t)})) = \frac{1}{T'} \sum_{t=t_1}^{T'} (S_{2\mathbf{L}_T}(\mathbf{f}^{(t)}) - S_2(\mathbf{f}^{(t)}))^2$ and $MSE(S_{2\mathbf{L}_A}(\mathbf{f}^{(t)})) = \frac{1}{T'} \sum_{t=t_1}^{T'} (S_{2\mathbf{L}_A}(\mathbf{f}^{(t)}) - S_2(\mathbf{f}^{(t)}))^2$.

Table 5.1 shows the average error and coding gain computed over each interval for $T_1 = 2$, $T_2 = 4$, and $T_3 = 6$. As observed, the proposed \mathbf{L}_T results in reduced error than \mathbf{L}_A with the network weights are high, whereas it is the opposite for smaller weights. Moreover, when the network connectivity is stronger (Interval 1), the error from both methods is inversely proportional to the duration of the interval. On the other hand, the average coding gain from both methods is close. Next, we assess the MSE of $S_2(\mathbf{f}^{(t)})$ for $\mathbf{f}_1^{(t)}$ and $\mathbf{f}_2^{(t)}$ computed over each interval, shown in Table 5.2 and Table 5.3, respectively. For both signals, this error is directly proportional to the error of the common Laplacians shown in Table 5.1. Together, these results suggest that the common subspace from \mathbf{L}_T preserves better the strong connectivity in the networks.

Table 5.3: MSE of the smoothness of $\mathbf{f}_2^{(t)}$ from \mathbf{L}_T and \mathbf{L}_A .

MSE				
	Interval 1		Interval 2	
T_1	\mathbf{L}_T	\mathbf{L}_A	\mathbf{L}_T	\mathbf{L}_A
2	104.016 ± 13.990	118.973 ± 15.990	94.671 ± 12.720	68.783 ± 9.250
4	87.660 ± 12.700	110.198 ± 15.980	132.627 ± 19.280	110.044 ± 16.000
6	50.121 ± 6.490	68.645 ± 8.830	128.681 ± 16.430	117.856 ± 15.030

5.3.2 Dynamic Functional Connectivity Networks

The proposed method is applied to EEG data obtained from the cognitive control experiment. The adjacency matrices for each subject S at time t , $\mathbf{A}_S^{(t)}$, are averaged over subjects to construct the dFCN $\hat{\mathbf{A}}^{(t)} = \frac{1}{T} \sum_{s=1}^S \mathbf{A}_s^{(t)}$. The Laplacian matrix $\mathbf{L}_A^{(t)}$ is then computed for this average adjacency. A 3-way tensor $\mathcal{L}^k \in \mathbb{R}^{N \times N \times T}$ is constructed for each time interval $k, = 1, \dots, 6$: [-25 ms, 0 ms], [0 ms, 25 ms], [25 ms, 50 ms], [50 ms, 75 ms], [75 ms, 100 ms], and [100 ms, 125 ms], where N corresponds to the number of electrodes, $N = 58$, and T to the total number of time points, $T = 14$ corresponding to the total number of time samples within each interval. The tensor \mathcal{L}^k is decomposed following (5.11) and the dGFT is computed for each interval k .

Figures 5.3 and 5.4 show the dGFT for each time interval from error and correct responses, respectively. As observed in Figure 5.3, the spectral energy from error responses is high around the ERN time interval [25 ms, 75 ms], specifically in the low frequencies. As shown in Figure 5.4, the spectral energy from correct responses remain mostly uniform except in intervals during the ERN where high energy is concentrated in the mid frequencies. The presence of high energy at low frequencies within the ERN interval in error responses reflects that during this time the graph signal is smooth with respect to the underlying network structure. On the other hand, high energy in higher frequencies suggests that the underlying network structure is less organized with respect to the graph signals.

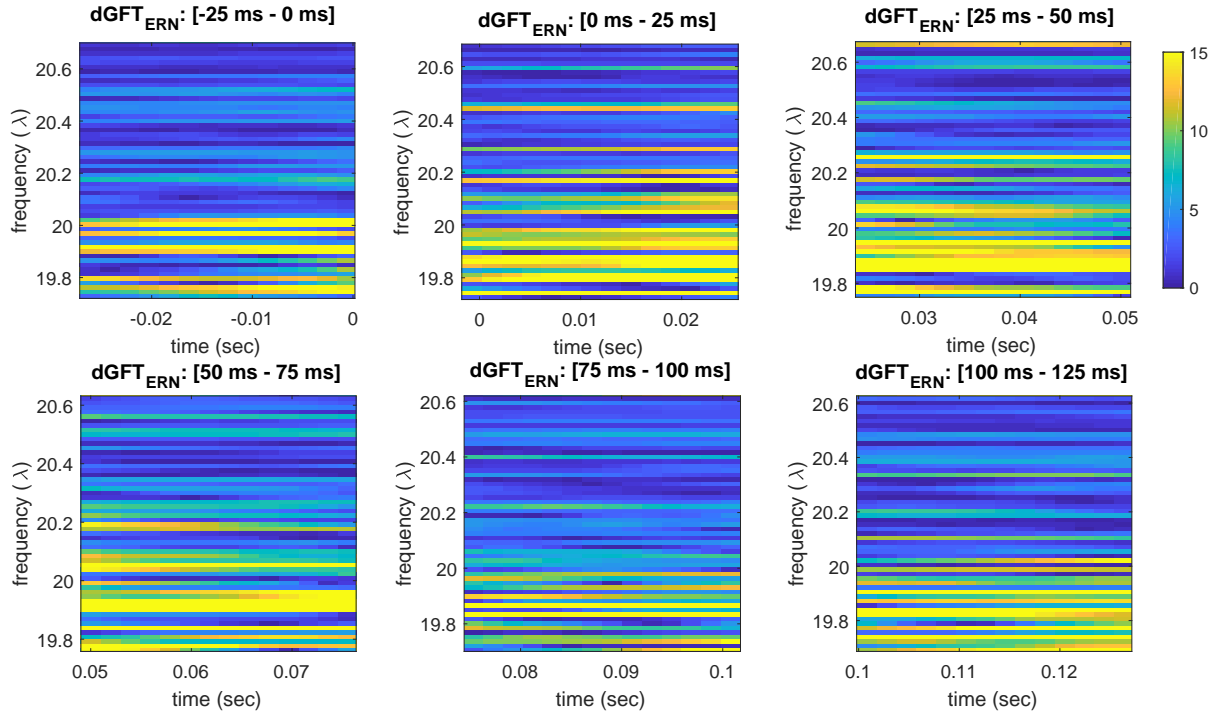


Figure 5.3: dGFT of the ERN dFCN over the interval [-25 ms, 125 ms].

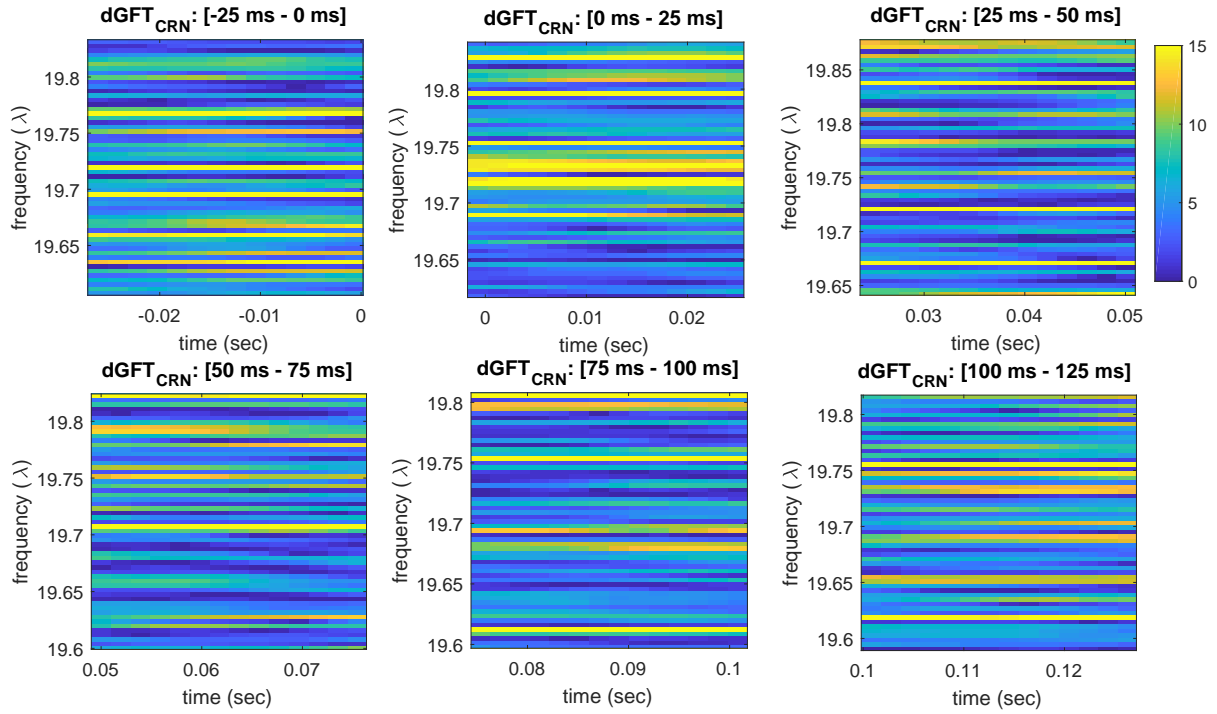


Figure 5.4: dGFT of the CRN dFCN over the interval [-25 ms, 125 ms].

For each interval we filter the graph signals $\mathbf{f}_{ERN}^{(t)}$ and $\mathbf{f}_{CRN}^{(t)}$ in the low, medium, and high frequency bands, where the low and middle bands are of equal length ($K = 20$) and the high frequency band length is $K = 18$ points. The signals are filtered in the frequency domain and then the iGFT is computed from the common basis for the corresponding window. Figures 5.5-5.10 show the topoplots corresponding to the graph signal in the middle of each interval, filtered in the three different frequency bands. For all time intervals, and for both response types, the highest energy is found in the low frequency band. It can be shown from these figures a high positive energy in the lateral and parietal regions during the ERN interval, with a strong negative component in the frontal-central regions. This is consistent with findings from previous works which relate lateral and central regions to be relevant during the ERN. On the other hand, CRN appears less organized, except in the [25 ms, 50 ms] interval.

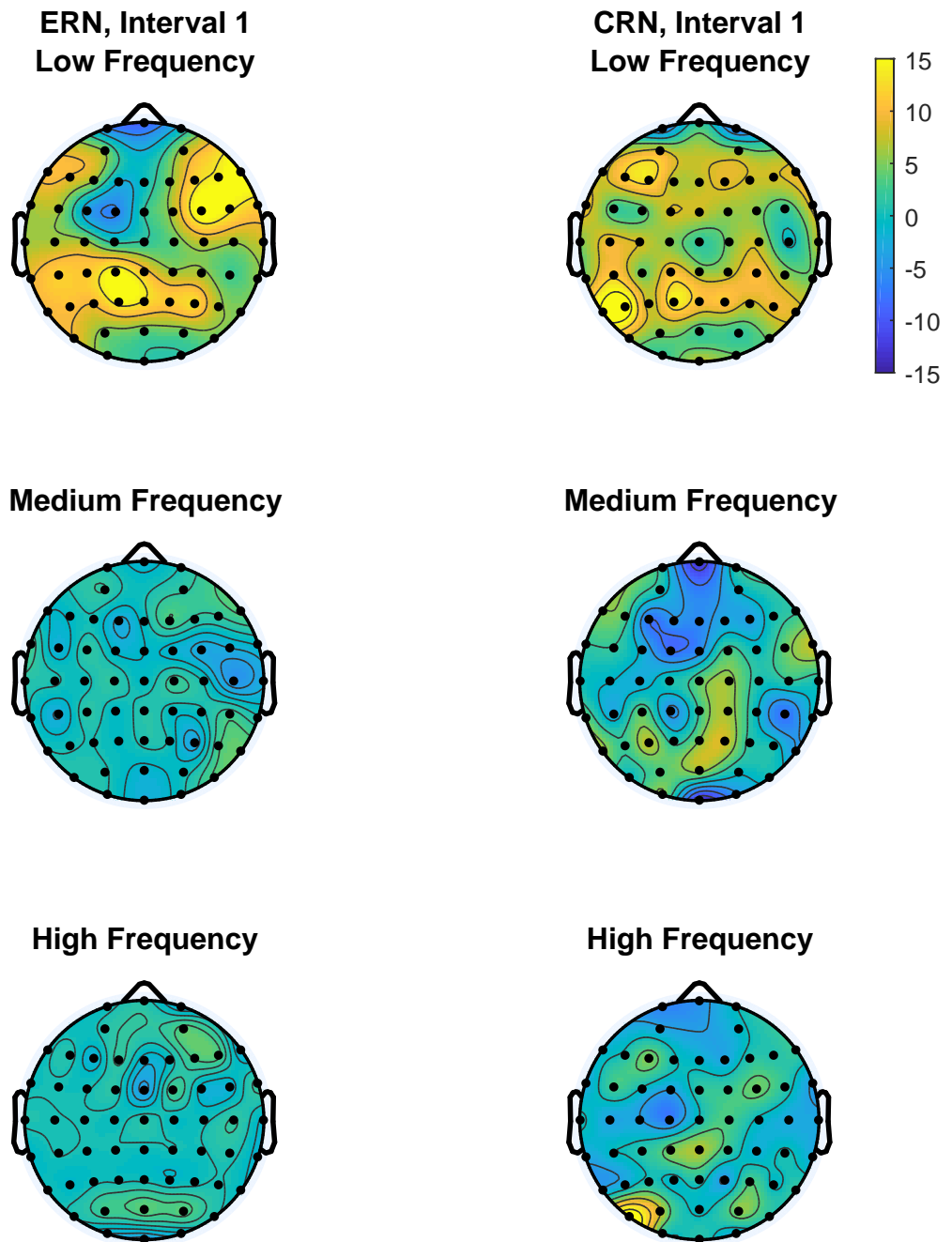


Figure 5.5: Topoplots from filtered signals in the low, medium and high graph frequency bands. Interval: [-25 ms, 0 ms].

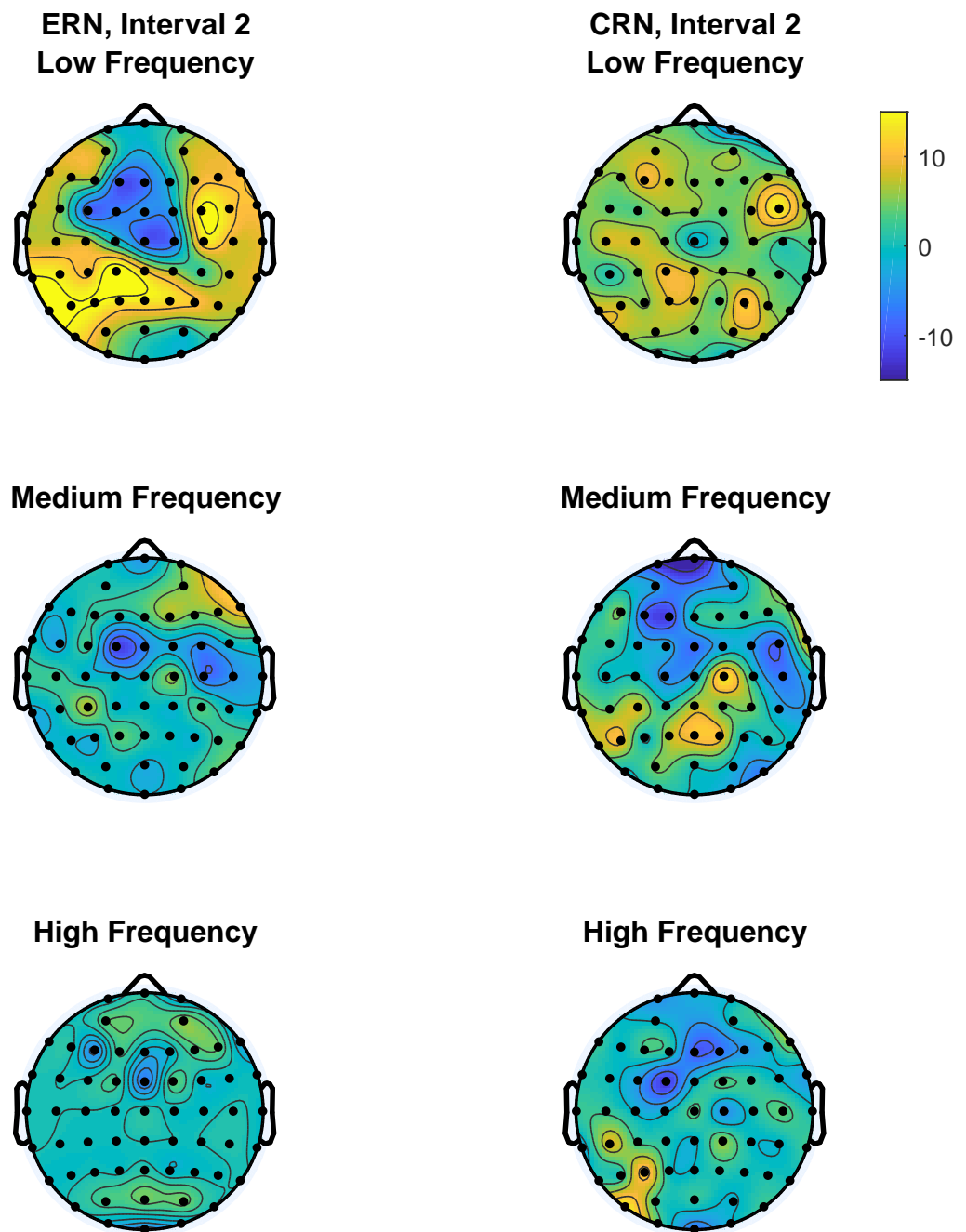


Figure 5.6: Topoplots from filtered signals in the low, medium and high graph frequency bands. Interval: [0 ms, 25 ms].

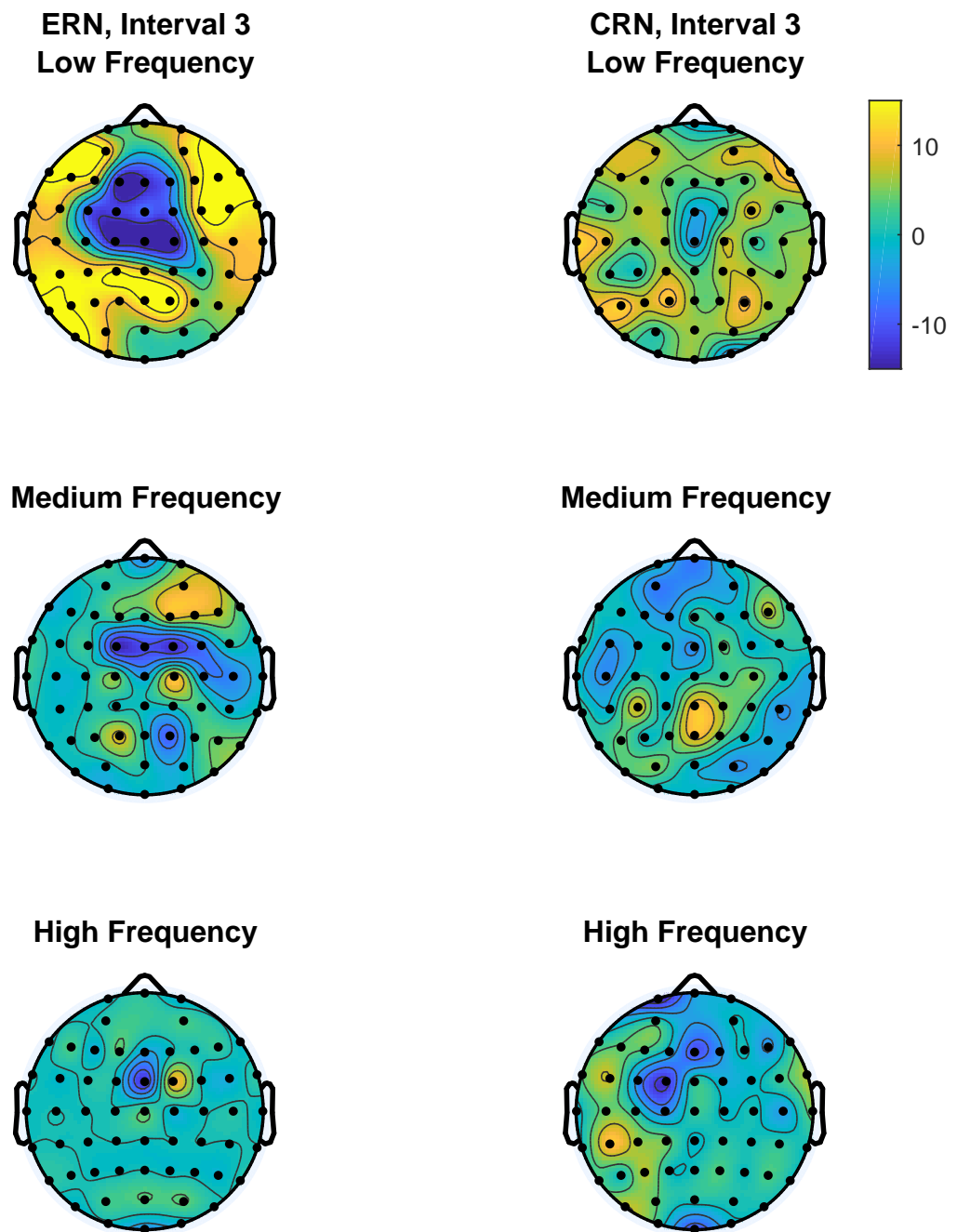


Figure 5.7: Topoplots from filtered signals in the low, medium and high graph frequency bands. Interval: [25 ms, 50 ms].

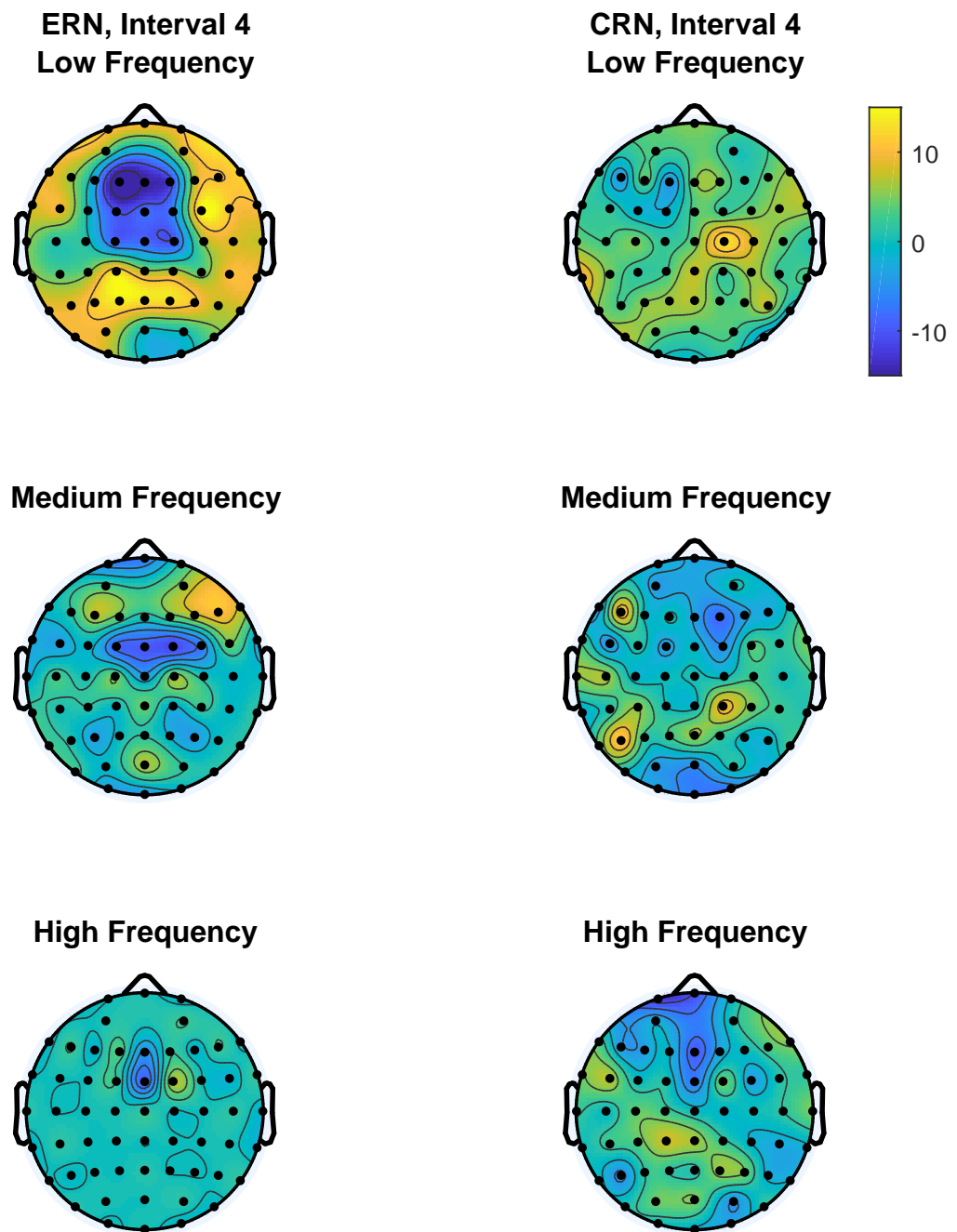


Figure 5.8: Topoplots from filtered signals in the low, medium and high graph frequency bands. Interval: [50 ms, 75 ms].

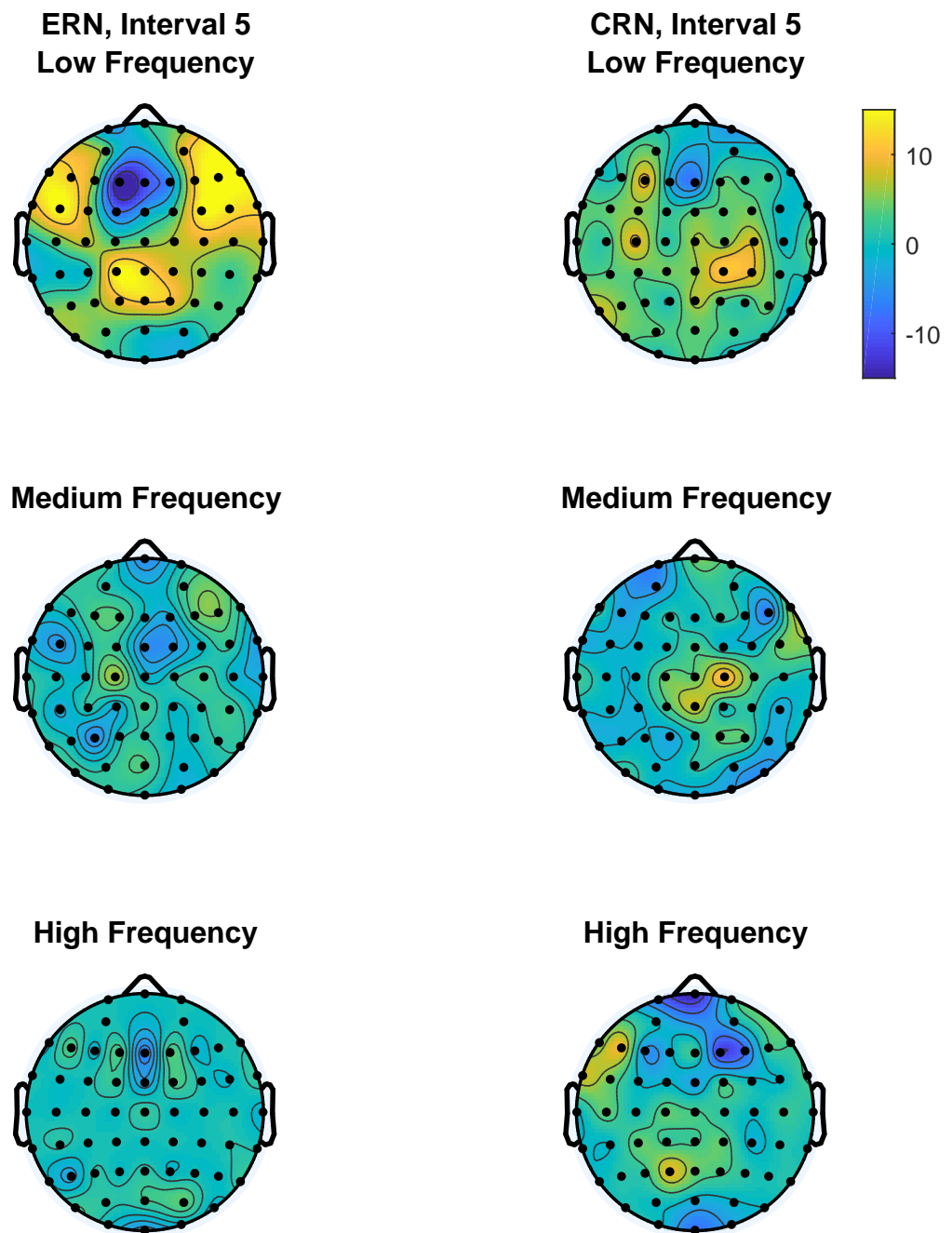


Figure 5.9: Topoplots from filtered signals in the low, medium and high graph frequency bands. Interval: [75 ms, 100 ms].

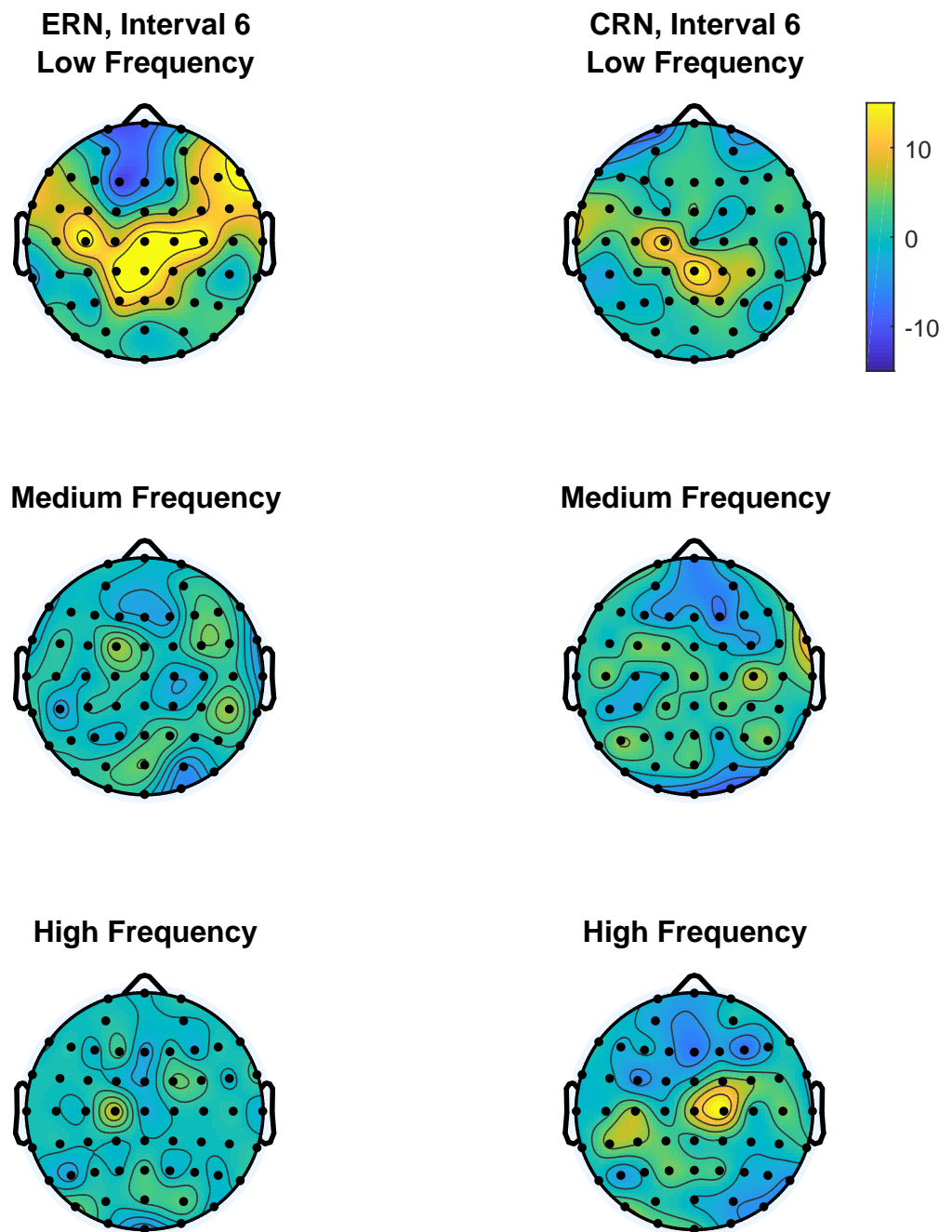


Figure 5.10: Topoplots from filtered signals in the low, medium and high graph frequency bands. Interval: [100 ms, 125 ms].

5.4 Conclusions

In this chapter, a dynamic graph Fourier transform based on the common basis obtained from the Tucker decomposition of the temporal network Laplacian tensor has been introduced to assess nonstationary networks. Simulations results show the effectiveness of this method under different scenarios. Furthermore, the proposed method was applied to EEG data from a cognitive control study to determine the brain regions that are highly involved in the ERN and to better understand the smoothness of the network during ERN. Future work will concentrate on extending the proposed method to account for both the common and the individual subspaces within each window, following the ideas from linked multiway tensor analysis. In addition, only stationary graph signals were considered. Future work will focus on extending non-stationary graph signals processing methods to account for nonstationarities in both the graph and the network domain.

Chapter 6

Conclusions and Future Work

In this thesis, we proposed a series of diverse techniques that aim to improve the quantification of multivariate functional connectivity in the brain. In Chapter 3, we addressed the problem of quantifying global synchrony, which cannot be addressed by traditional bivariate measures such as PLV. We introduced a novel measure of multivariate synchrony based on a hypertorus synchrony (HTS), which is equivalent to the Cartesian product of unit circles parameterized by phase differences. As PLV quantifies the variability of vectors in the unit circle parameterized by the phase difference between two oscillators, HTS quantifies the variability of unit vectors in the flat hypertorus parameterized by the phase difference between the current oscillator and the average phase of the other oscillators. Furthermore, by relying on the Cartesian product of circles, this definition ensures that the measure has good topographical sensitivity since it does not depend on the ordering of the phase differences.

This measure has a major impact on the study of functional connectivity since it allows for the quantification of global connectivity across different regions, not possible from bivariate PLV. Our results show that fronto-lateral and medial central regions exhibit greater functional integration during the ERN interval. Furthermore, the representation of functional integration as provided by HTS allows for correlating multivariate synchrony and behavioral measures of post-error adjustments, such as post-error accuracy (PEA) and post-error slowing (PES). Correlations between

HTS and PES compared to HTS and PEA show hemispherical topographical differences, suggesting that PES and PEA rely on different neural mechanisms. Specifically, it was shown that reduced PES is associated with increases in multivariate synchrony in frontal and parietal regions. This suggests that these regions are involved in the adaptivity between reduced PES and error-related neural activity. On the other hand, it was observed that PEA is associated with increased frontal synchronization, implying the integration of mPFC and LPFC in signaling and updating adaptive control mechanisms.

In Chapter 4, we introduce a method for transforming both weighted and binary networks into signals and applied this framework for the first time to the assessment of functional connectivity networks. In this work, we proposed to employ the resistance distance matrix as the distance matrix for classical multidimensional scaling (CMDS). The resistance distance is a valid Euclidean distance, and therefore guarantees the positive definiteness of the matrix for CMDS. Through this transformation, it is possible to overcome some of the drawbacks of graph theoretic measures. We showed that the signals obtained from the resistance matrix for binary networks follow results based on another previously proposed distance, and the signals from the resistance distance matrix contain additional information regarding the network structure. In addition, we showed that this transformation is robust to anomalies in the network and it is possible to reconstruct the original networks from the signals obtained from this transformation. Furthermore, transforming the networks into signals facilitates the computation of signal processing measures on graphs, such as information theoretic concepts. In this chapter, we presented a new graph entropy and divergence measure based on the signals from networks. The proposed methods apply the Shannon entropy and Kullback-Leibler divergence to the spectrum of the signals and are independent of any parameters opposed to current graph information theoretic measures. Finally, we proposed an event detection method for temporal networks based on the tensor decomposition of the magnitude spec-

tra of the signals over time, which is more sensitive to changes in the network structure. When applied to functional connectivity networks from a cognitive control experiment we observed the networks are well correlated with small-world networks, with networks from error responses being more correlated, following previous results. In addition, the proposed graph entropy measure suggests that functional connectivity networks from error responses during the ERN interval are more structurally organized than those from correct responses and divergence between error and correct networks increases during that interval.

Finally, in Chapter 5 we presented a dynamic graph Fourier transform (DGFT) which relies on tensor decomposition of the graph Laplacian over time for finding a common basis. This transform allows for the computation of the GFT for signals defined on time-varying networks, as it occurs in functional connectivity networks. Defining such a transform presents an alternative way of looking at functional connectivity networks, allowing for the study of the relationships between the responses recorded at the electrodes and its underlying network. Specifically, this brings a sense of smoothness and structural organization in the network. Our results demonstrate that error responses during the ERN interval result in a more structured and smooth network when compared to correct responses. Future work will focus on the computation of both common and individual subspace, which can be achieved from a linked multiway tensor decomposition. In addition, pre-processing for determining change points and defining the window length will be beneficial.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [1] P. J. Uhlhaas and W. Singer, “Neural synchrony in brain disorders: relevance for cognitive dysfunctions and pathophysiology,” *Neuron*, vol. 52, no. 1, pp. 155–168, 2006.
- [2] A. M. Bastos and J.-M. Schoffelen, “A tutorial review of functional connectivity analysis methods and their interpretational pitfalls,” *Frontiers in systems neuroscience*, vol. 9, 2015.
- [3] P. J. Uhlhaas and W. Singer, “Abnormal neural oscillations and synchrony in schizophrenia,” *Nature reviews neuroscience*, vol. 11, no. 2, pp. 100–113, 2010.
- [4] G. Buzsáki and A. Draguhn, “Neuronal oscillations in cortical networks,” *science*, vol. 304, no. 5679, pp. 1926–1929, 2004.
- [5] P. Fries, “Neuronal gamma-band synchronization as a fundamental process in cortical computation,” *Annual review of neuroscience*, vol. 32, pp. 209–224, 2009.
- [6] P. Lakatos, G. Karmos, A. D. Mehta, I. Ulbert, and C. E. Schroeder, “Entrainment of neuronal oscillations as a mechanism of attentional selection,” *science*, vol. 320, no. 5872, pp. 110–113, 2008.
- [7] P. J. Uhlhaas and W. Singer, “What do disturbances in neural synchrony tell us about autism?” *Biological psychiatry*, vol. 62, no. 3, pp. 190–191, 2007.
- [8] G. Tononi, “Functional segregation and integration in the nervous system: Theory and models,” in *Somesthesia and the Neurobiology of the Somatosensory Cortex*. Springer, 1996, pp. 409–418.
- [9] G. Tononi, O. Sporns, and G. M. Edelman, “A measure for brain complexity: relating functional segregation and integration in the nervous system,” *Proceedings of the National Academy of Sciences*, vol. 91, no. 11, pp. 5033–5037, 1994.
- [10] M. Rubinov and O. Sporns, “Complex network measures of brain connectivity: uses and interpretations,” *Neuroimage*, vol. 52, no. 3, pp. 1059–1069, 2010.
- [11] S. Baillet, J. C. Mosher, and R. M. Leahy, “Electromagnetic brain mapping,” *IEEE Signal Processing Magazine*, vol. 18, no. 6, pp. 14–30, 2001.

- [12] S. A. Huettel, A. W. Song, and G. McCarthy, *Functional magnetic resonance imaging*. Sinauer Associates Sunderland, 2004, vol. 1.
- [13] K. J. Friston, “Functional and effective connectivity: a review,” *Brain connectivity*, vol. 1, no. 1, pp. 13–36, 2011.
- [14] A. Rychwalska, “Understanding cognition through functional connectivity,” in *Complex Human Dynamics*. Springer, 2013, pp. 21–34.
- [15] A. Aertsen and H. Preissl, “Dynamics of activity and connectivity in physiological neuronal networks,” *Nonlinear dynamics and neuronal networks*, vol. 2, pp. 281–301, 1991.
- [16] L. Lee, L. M. Harrison, and A. Mechelli, “A report of the functional connectivity workshop, dusseldorf 2002,” *Neuroimage*, vol. 19, no. 2, pp. 457–465, 2003.
- [17] P. M. Matthews, N. Filippini, and G. Douaud, “Brain structural and functional connectivity and the progression of neuropathology in alzheimer’s disease,” *Journal of Alzheimer’s Disease*, vol. 33, no. s1, pp. S163–S172, 2013.
- [18] M.-E. Lynall, D. S. Bassett, R. Kerwin, P. J. McKenna, M. Kitzbichler, U. Muller, and E. Bullmore, “Functional connectivity and brain networks in schizophrenia,” *The Journal of Neuroscience*, vol. 30, no. 28, pp. 9477–9487, 2010.
- [19] J. F. Cavanagh and M. J. Frank, “Frontal theta as a mechanism for cognitive control,” *Trends in cognitive sciences*, vol. 18, no. 8, pp. 414–421, 2014.
- [20] C. S. Carter, T. S. Braver, D. M. Barch, M. M. Botvinick, D. Noll, and J. D. Cohen, “Anterior cingulate cortex, error detection, and the online monitoring of performance,” *Science*, vol. 280, no. 5364, pp. 747–749, 1998.
- [21] M. M. Botvinick, T. S. Braver, D. M. Barch, C. S. Carter, and J. D. Cohen, “Conflict monitoring and cognitive control,” *Psychological review*, vol. 108, no. 3, p. 624, 2001.
- [22] J. R. Hall, E. M. Bernat, and C. J. Patrick, “Externalizing psychopathology and the error-related negativity,” *Psychological Science*, vol. 18, no. 4, pp. 326–333, 2007.
- [23] K. R. Ridderinkhof, Y. de Vlugt, A. Bramlage, M. Spaan, M. Elton, J. Snel, and G. P. Band, “Alcohol consumption impairs detection of performance errors in mediofrontal cortex,” *Science*, vol. 298, no. 5601, pp. 2209–2211, 2002.

- [24] K. R. Ridderinkhof, M. Ullsperger, E. A. Crone, and S. Nieuwenhuis, "The role of the medial frontal cortex in cognitive control," *science*, vol. 306, no. 5695, pp. 443–447, 2004.
- [25] G. S. Alexopoulos, M. J. Hoptman, D. Kanellopoulos, C. F. Murphy, K. O. Lim, and F. M. Gunning, "Functional connectivity in the cognitive control network and the default mode network in late-life depression," *Journal of affective disorders*, vol. 139, no. 1, pp. 56–65, 2012.
- [26] L. Tian, T. Jiang, Y. Wang, Y. Zang, Y. He, M. Liang, M. Sui, Q. Cao, S. Hu, M. Peng *et al.*, "Altered resting-state functional connectivity patterns of anterior cingulate cortex in adolescents with attention deficit hyperactivity disorder," *Neuroscience letters*, vol. 400, no. 1, pp. 39–43, 2006.
- [27] J. Cavanagh, M. Cohen, and J. Allen, "Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring," *The Journal of Neuroscience*, vol. 29, no. 1, pp. 98–105, 2009.
- [28] E. Pereda, R. Quiroga, and J. Bhattacharya, "Nonlinear multivariate analysis of neurophysiological signals," *Progress in Neurobiology*, vol. 77, no. 1-2, pp. 1–37, 2005.
- [29] S. Aviyente, E. M. Bernat, W. S. Evans, and S. R. Sponheim, "A phase synchrony measure for quantifying dynamic functional integration in the brain," *Human brain mapping*, vol. 32, no. 1, pp. 80–93, 2011.
- [30] S. Aviyente and A. Mutlu, "A time-frequency-based approach to phase and phase synchrony estimation," *IEEE Transactions on Signal Processing*, vol. 59, no. 7, pp. 3086–3098, 2011.
- [31] S. Dimitriadis, N. Laskaris, and A. Tzelepi, "On the quantization of time-varying phase synchrony patterns into distinct functional connectivity microstates (fcμstates) in a multi-trial visual erp paradigm," *Brain topography*, vol. 26, no. 3, pp. 397–409, 2013.
- [32] J.-P. Lachaux, E. Rodriguez, J. Martinerie, F. J. Varela *et al.*, "Measuring phase synchrony in brain signals," *Human brain mapping*, vol. 8, no. 4, pp. 194–208, 1999.
- [33] M. Jalili, E. Barzegaran, and M. G. Knyazeva, "Synchronization of eeg: Bivariate and multivariate measures," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 2, pp. 212–221, 2014.
- [34] S. Aydore, D. Pantazis, and R. M. Leahy, "A note on the phase locking value and its properties," *Neuroimage*, vol. 74, pp. 231–244, 2013.

- [35] C. Stam and B. Van Dijk, “Synchronization likelihood: an unbiased measure of generalized synchronization in multivariate data sets,” *Physica D: Nonlinear Phenomena*, vol. 163, no. 3, pp. 236–251, 2002.
- [36] C. Carmeli, M. G. Knyazeva, G. M. Innocenti, and O. De Feo, “Assessment of EEG synchronization based on state-space analysis,” *Neuroimage*, vol. 25, no. 2, pp. 339–354, 2005.
- [37] A. Y. Mutlu and S. Aviyente, “Hyperspherical phase synchrony for quantifying multivariate phase synchronization,” in *2012 IEEE Statistical Signal Processing Workshop (SSP)*, 2012, pp. 888–891.
- [38] —, “Hyperspherical phase synchrony measure for quantifying global synchronization in the brain,” in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2013, pp. 1267–1271.
- [39] E. Bullmore and O. Sporns, “Complex brain networks: graph theoretical analysis of structural and functional systems,” *Nature Reviews Neuroscience*, vol. 10, no. 3, pp. 186–198, 2009.
- [40] D. S. Bassett and E. Bullmore, “Small-world brain networks,” *The neuroscientist*, vol. 12, no. 6, pp. 512–523, 2006.
- [41] M. P. van den Heuvel, C. J. Stam, M. Boersma, and H. H. 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.
- [42] A. Y. Mutlu and S. Aviyente, “Multivariate empirical mode decomposition for quantifying multivariate phase synchronization,” *EURASIP Journal on Advances in Signal Processing*, vol. 2011, pp. 1–13, 2011.
- [43] Y. Haraguchi, Y. Shimada, T. Ikeguchi, and K. Aihara, “Transformation from complex networks to time series using classical multidimensional scaling,” in *Artificial Neural Networks–ICANN 2009*. Springer, 2009, pp. 325–334.
- [44] Y. Shimada, T. Ikeguchi, and T. Shigehara, “From networks to time series,” *Physical review letters*, vol. 109, no. 15, p. 158701, 2012.
- [45] R. Hamon, P. Borgnat, P. Flandrin, and C. Robardet, “Duality between temporal networks and signals: Extraction of the temporal network structures,” *arXiv preprint arXiv:1505.03044*, 2015.

- [46] —, “From graphs to signals and back: Identification of network structures using spectral analysis,” 2015.
- [47] —, “From graphs to signals and back: Identification of network structures using spectral analysis,” *arXiv preprint arXiv:1502.04697*, 2015.
- [48] M. Dehmer and A. Mowshowitz, “A history of graph entropy measures,” *Information Sciences*, vol. 181, no. 1, pp. 57–78, 2011.
- [49] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, “The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains,” *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 83–98, 2013.
- [50] A. G. Mahyari and S. Aviyente, “Fourier transform for signals on dynamic graphs,” in *2014 48th Asilomar Conference on Signals, Systems and Computers*. IEEE, 2014, pp. 2001–2004.
- [51] W. Huang, L. Goldsberry, N. Wymbs, S. Grafton, D. Bassett, and A. Ribeiro, “Graph frequency analysis of brain signals,” *presented at the Graph Signal Processing Workshop, Philadelphia, United States, 2016*.
- [52] J. P. Lachaux, A. Lutz, D. Rudrauf, D. Cosmelli, M. Le Van Quyen, J. Martinerie, and F. Varela, “Estimating the time-course of coherence between single-trial brain signals: an introduction to wavelet coherence,” *Neurophysiologie Clinique/Clinical Neurophysiology*, vol. 32, no. 3, pp. 157–174, 2002.
- [53] M. D. Humphries and K. Gurney, “Network small-world-ness: a quantitative method for determining canonical network equivalence,” *PloS one*, vol. 3, no. 4, p. e0002051, 2008.
- [54] J. S. Moser, H. S. Schroder, C. Heeter, T. P. Moran, and Y.-H. Lee, “Mind your errors evidence for a neural mechanism linking growth mind-set to adaptive posterror adjustments,” *Psychological Science*, p. 0956797611419520, 2011.
- [55] B. A. Eriksen and C. W. Eriksen, “Effects of noise letters upon the identification of a target letter in a nonsearch task,” *Perception and psychophysics*, vol. 16, no. 1, pp. 143–149, 1974.
- [56] G. Gratton, M. G. Coles, and E. Donchin, “A new method for off-line removal of ocular artifact,” *Electroencephalography and clinical neurophysiology*, vol. 55, no. 4, pp. 468–484, 1983.

- [57] C. Liu, D. R. Weaver, S. H. Strogatz, and S. M. Reppert, “Cellular construction of a circadian clock: period determination in the suprachiasmatic nuclei,” *Cell*, vol. 91, no. 6, pp. 855–860, 1997.
- [58] F. Varela, J.-P. Lachaux, E. Rodriguez, and J. Martinerie, “The brainweb: phase synchronization and large-scale integration,” *Nature reviews neuroscience*, vol. 2, no. 4, pp. 229–239, 2001.
- [59] J. Buck, “Synchronous rhythmic flashing of fireflies. ii,” *Quarterly review of biology*, pp. 265–289, 1988.
- [60] K. Wiesenfeld, P. Colet, and S. H. Strogatz, “Frequency locking in josephson arrays: connection with the kuramoto model,” *Physical Review E*, vol. 57, no. 2, pp. 1563–1569, 1998.
- [61] S. H. Strogatz, D. M. Abrams, A. McRobie, B. Eckhardt, and E. Ott, “Theoretical mechanics: Crowd synchrony on the millennium bridge,” *Nature*, vol. 438, no. 7064, pp. 43–44, 2005.
- [62] N. J. Fliege and J. Wintermantel, “Complex digital oscillators and fsk modulators,” *IEEE Transactions on Signal Processing*, vol. 40, no. 2, pp. 333–342, 1992.
- [63] Y. Wang, F. Nunez, and F. J. Doyle, “Statistical analysis of the pulse-coupled synchronization strategy for wireless sensor networks,” *IEEE Transactions on Signal Processing*, vol. 61, no. 21, pp. 5193–5204, 2013.
- [64] Y. Wang, F. Nunez, and F. J. Doyle III, “Energy-efficient pulse-coupled synchronization strategy design for wireless sensor networks through reduced idle listening,” *IEEE Transactions on Signal Processing*, vol. 60, no. 10, pp. 5293–5306, 2012.
- [65] Y. Wang and F. J. Doyle, “Optimal phase response functions for fast pulse-coupled synchronization in wireless sensor networks,” *IEEE Transactions on Signal Processing*, vol. 60, no. 10, pp. 5583–5588, 2012.
- [66] M. Rosenblum, A. Pikovsky, J. Kurths, C. Schäfer, and P. A. Tass, “Phase synchronization: from theory to data analysis,” *Handbook of biological physics*, vol. 4, pp. 279–321, 2001.
- [67] A. Pikovsky, M. Rosenblum, and J. Kurths, “Synchronization: A universal concept in non-linear systems,” *Cambridge Nonlinear Science Series*, vol. 12, 2001.
- [68] S. Boccaletti, J. Kurths, G. Osipov, D. Valladares, and C. Zhou, “The synchronization of chaotic systems,” *Physics Reports*, vol. 366, no. 1, pp. 1–101, 2002.

- [69] M. Le Van Quyen, J. Foucher, J. Lachaux, E. Rodriguez, A. Lutz, J. Martinerie, and F. Varela, "Comparison of hilbert transform and wavelet methods for the analysis of neuronal synchrony," *Journal of neuroscience methods*, vol. 111, no. 2, pp. 83–98, 2001.
- [70] T. M. Rutkowski, D. P. Mandic, A. Cichocki, and A. W. Przybyszewski, "EMD approach to multichannel EEG data the amplitude and phase components clustering analysis," *Journal of Circuits, Systems, and Computers*, vol. 19, no. 01, pp. 215–229, 2010.
- [71] A. Ahrabian and D. P. Mandic, "Estimation of phase synchrony using the synchrosqueezing transform," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 759–763.
- [72] A. Omidvarnia, G. Azemi, P. B. Colditz, and B. Boashash, "A time–frequency based approach for generalized phase synchrony assessment in nonstationary multivariate signals," *Digital Signal Processing*, vol. 23, no. 3, pp. 780–790, 2013.
- [73] A. Ahrabian, C. C. Took, and D. P. Mandic, "Algorithmic trading using phase synchronization," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 4, pp. 399–404, 2012.
- [74] S. Aviyente, E. Bernat, W. Evans, and S. Sponheim, "A phase synchrony measure for quantifying dynamic functional integration in the brain," *Human brain mapping*, vol. 32, no. 1, pp. 80–93, 2011.
- [75] R. Q. Quiroga, A. Kraskov, T. Kreuz, and P. Grassberger, "Performance of different synchronization measures in real data: a case study on electroencephalographic signals," *Physical Review E*, vol. 65, no. 4, p. 041903, 2002.
- [76] D. Cui, X. Liu, Y. Wan, and X. Li, "Estimation of genuine and random synchronization in multivariate neural series," *Neural Networks*, vol. 23, no. 6, pp. 698–704, 2010.
- [77] T. Koenig, D. Lehmann, N. Saito, T. Kuginuki, T. Kinoshita, and M. Koukkou, "Decreased functional connectivity of EEG theta-frequency activity in first-episode, neuroleptic-naïve patients with schizophrenia: preliminary results," *Schizophrenia research*, vol. 50, no. 1, pp. 55–60, 2001.
- [78] M. Jalili, S. Lavoie, P. Deppen, R. Meuli, K. Q. Do, M. Cuénod, M. Hasler, O. De Feo, and M. G. Knyazeva, "Dysconnection topography in schizophrenia revealed with state-space analysis of EEG," *PLoS One*, vol. 2, no. 10, p. e1059, 2007.

- [79] M. Jalili, E. Barzegaran, and M. Knyazeva, "Synchronization of EEG: Bivariate and multivariate measures," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 2, pp. 212–221, 2014.
- [80] M. G. Knyazeva, M. Jalili, A. Brioschi, I. Bourquin, E. Fornari, M. Hasler, R. Meuli, P. Maeder, and J. Ghika, "Topography of EEG multivariate phase synchronization in early alzheimer's disease," *Neurobiology of Aging*, vol. 31, no. 7, pp. 1132–1144, 2010.
- [81] D. Rudrauf, A. Douiri, C. Kovach, J. Lachaux, D. Cosmelli, M. Chavez, C. Adam, B. Renault, J. Martinerie, and M. Le Van Quyen, "Frequency flows and the time-frequency dynamics of multivariate phase synchronization in brain signals," *Neuroimage*, vol. 31, no. 1, pp. 209–227, 2006.
- [82] C. Allefeld and J. Kurths, "An approach to multivariate phase synchronization analysis and its application to event-related potentials," *International Journal of Bifurcation and Chaos*, vol. 14, no. 2, pp. 417–426, 2004.
- [83] M. J. Richardson, R. L. Garcia, T. D. Frank, M. Gergor, and K. L. Marsh, "Measuring group synchrony: a cluster-phase method for analyzing multivariate movement time-series," *Frontiers in physiology*, vol. 3, 2012.
- [84] C. Allefeld, M. Müller, and J. Kurths, "Eigenvalue decomposition as a generalized synchronization cluster analysis," *International Journal of Bifurcation and Chaos*, vol. 17, pp. 3493–3497, 2007.
- [85] A. S. Fine, D. P. Nicholls, and D. J. Mogul, "Assessing instantaneous synchrony of nonlinear nonstationary oscillators in the brain," *Journal of neuroscience methods*, vol. 186, no. 1, pp. 42–51, 2010.
- [86] C. Allefeld and S. Bialonski, "Detecting synchronization clusters in multivariate time series via coarse-graining of Markov chains," *Physical Review E*, vol. 76, no. 6, pp. 66 207–66 215, 2007.
- [87] N. Saito, T. Kuginuki, T. Yagyu, T. Kinoshita, T. Koenig, R. D. Pascual-Marqui, K. Kochi, J. Wackermann, and D. Lehmann, "Global, regional, and local measures of complexity of multichannel electroencephalography in acute, neuroleptic-naive, first-break schizophrenics," *Biological psychiatry*, vol. 43, no. 11, pp. 794–802, 1998.
- [88] J. Dauwels, F. Vialatte, 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.

- [89] M. Al-Khassaweneh, M. Villafañe-Delgado, A. Y. Mutlu, and S. Aviyente, “A measure of multivariate phase synchrony using hyperdimensional geometry,” *IEEE Transactions on Signal Processing*, vol. 64, no. 11, pp. 2774–2787, 2016.
- [90] N. I. Fisher, *Statistical analysis of circular data*. Cambridge University Press, 1995.
- [91] J. M. Lee, *Riemannian manifolds: an introduction to curvature*. Springer Science & Business Media, 2006, vol. 176.
- [92] J. Jost, *Riemannian geometry and geometric analysis*. Springer Science & Business Media, 2008.
- [93] M. L. Oristaglio and B. R. Spies, *Three-dimensional electromagnetics*. SEG Books, 1999, no. 7.
- [94] P. Renteln, *Manifolds, Tensors, and Forms: An Introduction for Mathematicians and Physicists*. Cambridge University Press, 2013.
- [95] S. Mallat, *A wavelet tour of signal processing*. Academic press, 1999.
- [96] D. Henderson, “Elementary functions: Algorithms and implementation,” *Mathematics and Computer Education*, vol. 34, no. 1, p. 94, 2000.
- [97] J. Demmel, I. Dumitriu, and O. Holtz, “Fast linear algebra is stable,” *Numerische Mathematik*, vol. 108, no. 1, pp. 59–91, 2007.
- [98] I. N. Bronštejn and K. A. Semendjaev, *Handbook of mathematics*. Springer, 2013.
- [99] K. Mardia, “Statistics of directional data: probability and mathematical statistics. 357 pp,” 1972.
- [100] M. Vinck, M. van Wingerden, T. Womelsdorf, P. Fries, and C. M. Pennartz, “The pairwise phase consistency: a bias-free measure of rhythmic neuronal synchronization,” *Neuroimage*, vol. 51, no. 1, pp. 112–122, 2010.
- [101] Y. Kuramoto, “Self-entrainment of a population of coupled non-linear oscillators,” in *International symposium on mathematical problems in theoretical physics*, 1975, pp. 420–422.
- [102] S. H. Strogatz, “From kuramoto to crawford: exploring the onset of synchronization in populations of coupled oscillators,” *Physica D: Nonlinear Phenomena*, vol. 143, no. 1, pp. 1–20, 2000.

- [103] A. Takamatsu, T. Fujii, and I. Endo, “Time delay effect in a living coupled oscillator system with the plasmodium of *Physarum polycephalum*,” *Physical review letters*, vol. 85, no. 9, pp. 2026–2029, 2000.
- [104] C. J. 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.
- [105] S. H. Strogatz, “Exploring complex networks,” *Nature*, vol. 410, no. 6825, pp. 268–276, 2001.
- [106] B. Schelter, M. Winterhalder, R. Dahlhaus, J. Kurths, and J. Timmer, “Partial phase synchronization for multivariate synchronizing systems,” *Physical Review Letters*, vol. 96, no. 20, p. 208103, 2006.
- [107] B. Veeramani, K. Narayanan, A. Prasad, L. D. Iasemidis, A. S. Spanias, and K. Tsakalis, “Measuring the direction and the strength of coupling in nonlinear systems-a modeling approach in the state space,” *IEEE Signal Processing Letters*, vol. 11, no. 7, pp. 617–620, 2004.
- [108] O. De Feo, G. M. Maggio, and M. P. Kennedy, “The colpitts oscillator: Families of periodic solutions and their bifurcations,” *International journal of bifurcation and chaos*, vol. 10, no. 05, pp. 935–958, 2000.
- [109] T. P. Moran, E. M. Bernat, S. Aviyente, H. S. Schroder, and J. S. Moser, “Sending mixed signals: Worry is associated with enhanced initial error processing but reduced call for subsequent cognitive control,” *Social cognitive and affective neuroscience*, p. nsv046, 2015.
- [110] H. S. Schroder and J. S. Moser, “Improving the study of error monitoring with consideration of behavioral performance measures,” *Frontiers in human neuroscience*, vol. 8, 2014.
- [111] G. A. Pavlopoulos, M. Secier, C. N. Moschopoulos, T. G. Soldatos, S. Kossida, J. Aerts, R. Schneider, and P. G. Bagos, “Using graph theory to analyze biological networks,” *Bio-Data mining*, vol. 4, no. 1, p. 1, 2011.
- [112] K. Park, “The internet as a complex system,” *The Internet as a Large-Scale Complex System. Santa Fe Institute Studies on the Sciences of Complexity*, Oxford University Press, Oxford, 2005.
- [113] P. J. Carrington, J. Scott, and S. Wasserman, *Models and methods in social network analysis*. Cambridge university press, 2005, vol. 28.

- [114] Q. K. Telesford, K. E. Joyce, S. Hayasaka, J. H. Burdette, and P. J. Laurienti, “The ubiquity of small-world networks,” *Brain connectivity*, vol. 1, no. 5, pp. 367–375, 2011.
- [115] R. Albert and A.-L. Barabási, “Statistical mechanics of complex networks,” *Reviews of modern physics*, vol. 74, no. 1, p. 47, 2002.
- [116] M. Bolanos, E. M. Bernat, B. He, and S. Aviyente, “A weighted small world network measure for assessing functional connectivity,” *Journal of neuroscience methods*, vol. 212, no. 1, pp. 133–142, 2013.
- [117] M. Rubinov and O. Sporns, “Complex network measures of brain connectivity: Uses and interpretations,” *NeuroImage*, vol. 52, no. 3, pp. 1059 – 1069, 2010.
- [118] M. Zanin, “On alternative formulations of the small-world metric in complex networks,” *arXiv preprint arXiv:1505.03689*, 2015.
- [119] D. Papo, M. Zanin, J. A. Pineda-Pardo, S. Boccaletti, and J. M. Buldú, “Functional brain networks: great expectations, hard times and the big leap forward,” *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, vol. 369, no. 1653, p. 20130525, 2014.
- [120] E. Estrada and N. Hatano, “Communicability in complex networks,” *Physical Review E*, vol. 77, no. 3, p. 036111, 2008.
- [121] T. Weng, Y. Zhao, M. Small, and D. D. Huang, “Time-series analysis of networks: Exploring the structure with random walks,” *Physical Review E*, vol. 90, no. 2, p. 022804, 2014.
- [122] B. Girault, P. Gonçalves, E. Fleury, and A. S. Mor, “Semi-supervised learning for graph to signal mapping: a graph signal wiener filter interpretation,” in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2014, pp. 1115–1119.
- [123] X. Li, X. Liu, and C. K. Tse, “Recent advances in bridging time series and complex networks,” in *Circuits and Systems (ISCAS), 2013 IEEE International Symposium on*. IEEE, 2013, pp. 2505–2508.
- [124] R. Hamon, P. Borgnat, P. Flandrin, and C. Robardet, “Nonnegative matrix factorization to find features in temporal networks,” in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 1065–1069.
- [125] —, “Extraction of temporal network structures from graph-based signals,” vol. 2, no. 2, pp. 215–226, 2016.

- [126] D. J. Klein and M. Randić, “Resistance distance,” *Journal of Mathematical Chemistry*, vol. 12, no. 1, pp. 81–95, 1993.
- [127] R. Bapat, “Resistance matrix of a weighted graph,” *Communications in Mathematical and in Computer Chemistry/MATCH*, vol. 50, pp. 73–82, 2004.
- [128] J. R. Sato, D. Y. Takahashi, M. Q. Hoexter, K. B. Massirer, and A. Fujita, “Measuring network’s entropy in adhd: A new approach to investigate neuropsychiatric disorders,” *NeuroImage*, vol. 77, pp. 44–51, 2013.
- [129] D. Y. Takahashi, J. R. Sato, C. E. Ferreira, and A. Fujita, “Discriminating different classes of biological networks by analyzing the graphs spectra distribution,” *PloS one*, vol. 7, no. 12, p. e49949, 2012.
- [130] M. Dehmer, L. Sivakumar, and K. Varmuza, “Uniquely discriminating molecular structures using novel eigenvaluebased descriptors,” *Match-Communications in Mathematical and Computer Chemistry*, vol. 67, no. 1, p. 147, 2012.
- [131] M. Dehmer, “Information processing in complex networks: Graph entropy and information functionals,” *Applied Mathematics and Computation*, vol. 201, no. 1, pp. 82–94, 2008.
- [132] P. Zhu and R. C. Wilson, “A study of graph spectra for comparing graphs.” in *BMVC*, 2005.
- [133] N. Rashevsky, “Life, information theory, and topology,” *The bulletin of mathematical biophysics*, vol. 17, no. 3, pp. 229–235, 1955.
- [134] A. Mowshowitz, “Entropy and the complexity of graphs: I. an index of the relative complexity of a graph,” *The bulletin of mathematical biophysics*, vol. 30, no. 1, pp. 175–204, 1968.
- [135] M. Dehmer, S. Borgert, and F. Emmert-Streib, “Network classes and graph complexity measures,” in *Proceedings of the 2008 First International Conference on Complexity and Intelligence of the Artificial and Natural Complex Systems. Medical Applications of the Complex Systems. Biomedical Computing*, 2008, pp. 77–84.
- [136] S. Ranshous, S. Shen, D. Koutra, S. Harenberg, C. Faloutsos, and N. F. Samatova, “Anomaly detection in dynamic networks: a survey,” *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 7, no. 3, pp. 223–247, 2015.
- [137] S. I. Dimitriadis, N. A. Laskaris, V. Tsirka, M. Vourkas, S. Micheloyannis, and S. Fotopoulos, “Tracking brain dynamics via time-dependent network analysis,” *Journal of neuroscience methods*, vol. 193, no. 1, pp. 145–155, 2010.

- [138] L. Gauvin, A. Panisson, and C. Cattuto, “Detecting the community structure and activity patterns of temporal networks: a non-negative tensor factorization approach,” *PloS one*, vol. 9, no. 1, p. e86028, 2014.
- [139] A. Cichocki, D. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, and H. A. Phan, “Tensor decompositions for signal processing applications: From two-way to multiway component analysis,” *Signal Processing Magazine, IEEE*, vol. 32, no. 2, pp. 145–163, 2015.
- [140] O. Sporns and J. D. Zwi, “The small world of the cerebral cortex,” *Neuroinformatics*, vol. 2, no. 2, pp. 145–162, 2004.
- [141] C. Ye, R. C. Wilson, and E. R. Hancock, “An entropic edge assortativity measure,” in *International Workshop on Graph-Based Representations in Pattern Recognition*. Springer, 2015, pp. 23–33.
- [142] S. L. Braunstein, S. Ghosh, and S. Severini, “The laplacian of a graph as a density matrix: a basic combinatorial approach to separability of mixed states,” *Annals of Combinatorics*, vol. 10, no. 3, pp. 291–317, 2006.
- [143] A. Li and Y. Pan, “Structural information and dynamical complexity of networks,” *IEEE Transactions on Information Theory*, vol. 62, no. 6, pp. 3290–3339, 2016.
- [144] F. Rossi, “Visualization methods for metric studies,” in *Proceedings of the International Workshop on Webometrics, Informetrics and Scientometrics*, 2006, pp. 356–366.
- [145] A. Ghosh, S. Boyd, and A. Saberi, “Minimizing effective resistance of a graph,” *SIAM review*, vol. 50, no. 1, pp. 37–66, 2008.
- [146] W. Ellens, F. Spieksma, P. Van Mieghem, A. Jamakovic, and R. Kooij, “Effective graph resistance,” *Linear algebra and its applications*, vol. 435, no. 10, pp. 2491–2506, 2011.
- [147] R. B. Bapat, *Graphs and matrices*. Springer, 2010.
- [148] W. Ellens and R. E. Kooij, “Graph measures and network robustness,” *arXiv preprint arXiv:1311.5064*, 2013.
- [149] A. Tizghadam and A. Leon-Garcia, “On iterative calculation of moore-penrose laplacian and resistance distance.”
- [150] R. Balaji and R. Bapat, “On euclidean distance matrices,” *Linear algebra and its applications*, vol. 424, no. 1, pp. 108–117, 2007.

- [151] P. Honeine, “An eigenanalysis of data centering in machine learning,” *arXiv preprint arXiv:1407.2904*, 2014.
- [152] H. Karner, J. Schneid, and C. W. Ueberhuber, “Spectral decomposition of real circulant matrices,” *Linear Algebra and Its Applications*, vol. 367, pp. 301–311, 2003.
- [153] M. Villafañe-Delgado and S. Aviyente, “Graph information theoretic measures on functional connectivity networks based on graph-to-signal transform,” in *Signal and Information Processing (GlobalSIP), 2016 IEEE Global Conference on*. IEEE, 2016, pp. 1137–1141.
- [154] T. G. Kolda and B. W. Bader, “Tensor decompositions and applications,” *SIAM review*, vol. 51, no. 3, pp. 455–500, 2009.
- [155] B. W. Bader and T. G. Kolda, “Matlab tensor toolbox version 2.6,” Available online, February 2015. [Online]. Available: <http://www.sandia.gov/tgkolda/TensorToolbox/>
- [156] —, “Algorithm 862: MATLAB tensor classes for fast algorithm prototyping,” *ACM Transactions on Mathematical Software*, vol. 32, no. 4, pp. 635–653, December 2006.
- [157] D. D. Lee and H. S. Seung, “Algorithms for non-negative matrix factorization,” in *Advances in neural information processing systems*, 2001, pp. 556–562.
- [158] R. Bro and H. A. Kiers, “A new efficient method for determining the number of components in parafac models,” *Journal of chemometrics*, vol. 17, no. 5, pp. 274–286, 2003.
- [159] A. Ozdemir, E. M. Bernat, and S. Aviyente, “Recursive tensor subspace tracking for dynamic brain network analysis,” *IEEE Transactions on Signal and Information Processing over Networks*, 2017.
- [160] C. Danielmeier, T. Eichele, B. U. Forstmann, M. Tittgemeyer, and M. Ullsperger, “Posterior medial frontal cortex activity predicts post-error adaptations in task-related visual and motor areas,” *The Journal of Neuroscience*, vol. 31, no. 5, pp. 1780–1789, 2011.
- [161] A. Navarro-Cebrian, R. T. Knight, and A. S. Kayser, “Frontal monitoring and parietal evidence: Mechanisms of error correction,” *Journal of cognitive neuroscience*, 2016.
- [162] D. K. Hammond, P. Vandergheynst, and R. Gribonval, “Wavelets on graphs via spectral graph theory,” *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129–150, 2011.

- [163] N. Saito, “Data analysis and representation on a general domain using eigenfunctions of laplacian,” *Applied and Computational Harmonic Analysis*, vol. 25, no. 1, pp. 68–97, 2008.
- [164] W. Huang, L. Goldsberry, N. F. Wymbs, S. T. Grafton, D. S. Bassett, and A. Ribeiro, “Graph frequency analysis of brain signals,” *arXiv preprint arXiv:1512.00037*, 2015.
- [165] A. Sandryhaila and J. M. Moura, “Discrete signal processing on graphs: Frequency analysis,” *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3042–3054, 2014.
- [166] D. I. Shuman, B. Ricaud, and P. Vandergheynst, “Vertex-frequency analysis on graphs,” *Applied and Computational Harmonic Analysis*, vol. 40, no. 2, pp. 260–291, 2016.
- [167] A. Loukas and D. Foucard, “Frequency analysis of temporal graph signals,” *arXiv preprint arXiv:1602.04434*, 2016.
- [168] F. Grassi, N. Perraudin, and B. Ricaud, “Tracking time-vertex propagation using dynamic graph wavelets,” *arXiv preprint arXiv:1606.06653*, 2016.
- [169] R. M. Hutchison, T. Womelsdorf, E. A. Allen, P. A. Bandettini, V. D. Calhoun, M. Corbetta, S. Della Penna, J. H. Duyn, G. H. Glover, J. Gonzalez-Castillo *et al.*, “Dynamic functional connectivity: promise, issues, and interpretations,” *Neuroimage*, vol. 80, pp. 360–378, 2013.
- [170] X. Liu, S. Ji, W. Glänzel, and B. De Moor, “Multiview partitioning via tensor methods,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 5, pp. 1056–1069, 2013.
- [171] E. Acar and B. Yener, “Unsupervised multiway data analysis: A literature survey,” *IEEE transactions on knowledge and data engineering*, vol. 21, no. 1, pp. 6–20, 2009.
- [172] W. Tang, Z. Lu, and I. S. Dhillon, “Clustering with multiple graphs,” in *2009 Ninth IEEE International Conference on Data Mining*. IEEE, 2009, pp. 1016–1021.
- [173] A. Argyriou, M. Herbster, and M. Pontil, “Combining graph laplacians for semi-supervised learning,” in *Advances in Neural Information Processing Systems*, 2005, pp. 67–74.
- [174] H. E. Egilmez, A. Ortega, O. G. Guleryuz, J. Ehmann, and S. Yea, “An optimization framework for combining multiple graphs,” in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2016, pp. 4114–4118.