Mutual Information Derived Functional Connectivity of the Electroencephalogram (EEG)

by

Pamela Wen-Hsin Lee

B.A.Sc., The University of British Columbia, 2005

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

Master of Applied Science

in

The Faculty of Graduate Studies

(Electrical and Computer Engineering)

The University of British Columbia

December, 2007

Abstract

Monitoring the functional connectivity between brain networks is becoming increasingly important in elucidating brain functionality in normal and disease states. Current methods of detecting networks in the recorded EEG such as correlation and coherence are limited by the fact that they assume stationarity of the relationship between channels, and rely on linear dependencies. Here we utilize mutual information (MI) as the metric for determining nonlinear statistical dependencies between electroencephalographic (EEG) channels. Previous work investigating MI between EEG channels in subjects with widespread diseases of the cerebral cortex had subjects simply rest quietly with their eyes closed. In motor disorders such as Parkinson's disease (PD), abnormalities are only expected during performance of motor tasks, but this makes the assumption of stationarity of relationships between EEG channels untenable. We therefore propose a novel EEG segmentation method based on the temporal dynamics of the cross-spectrogram of the computed Independent Components (ICs). After suitable thresholding of the MI values between channels in the temporally segmented EEG, graphical theoretical analysis approaches are applied to the derived networks. The method was applied to EEG data recorded from six normal subjects and seven PD subjects on and off medication performing a motor task involving either their right hand only or both hands simultaneously. One-way analysis of variance (ANOVA) tests demonstrated statistically significant difference between subject groups. This proposed segmentation/MI network method appears to be a promising approach for EEG analysis.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank everyone who has helped and inspired me during my master study.

I especially want to convey my appreciation to my supervisor, Dr. Jane Wang, whose patient guidance, persistent encouragement and deep insight in the research area have helped me throughout the course of my thesis research.

Dr. Martin McKeown deserves a special thanks for his guidance during my research. He was always accessible and willing to help. His knowledge of Parkinson's disease is second to none and he provided me with many insightful comments and suggestions.

I am grateful for my thesis examination committee members who monitored my work and took effort in reading and providing me with valuable comments. As well, I would like to thank Samantha Palmer for collecting the data.

My deepest gratitude goes to my family for their endless love, constant support, and immense encouragement over the years. It is their constant prayers that allowed me to go on.

Last but not least, I would like to thank God for being my strength throughout my life. It is by His grace and mercy that I am able to do all things. May His name be exalted, honored, and glorified.

This work was supported by the NSERC grant (JW) and the Collaborative Health Research Projects (CHRPJ) grant.

Co-Authorship Statement

The author, Pamela W. H. Lee, has performed the research and data analyses and also prepared the manuscript and conference proceedings in this thesis. The research work was initiated by Dr. Z. Jane Wang and Dr. Martin J. McKeown. They have helped greatly on revisions of the manuscript and conference proceedings. The data are collected by Samantha Palmer.

Chapter 1

Introduction

1.1 Background and Motivation

Extensive effort has been put forth to understand how the human brain works. Intense research has been conducted at the molecular, cellular, and ultimately macroscopic levels. With recent advances in electronics and signal processing, a variety of macroscopic measures can now be widely used in both clinical and research settings to examine the function and structure of the human brain. These technologies include electroencephalography (EEG), magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), and positron emission tomography (PET). EEG measures the electrical activity of the brain and represents a summation of post-synaptic potentials from a large number of neurons [1]. EEG recording, which involves placing electrodes on the scalp, is a non-invasive method for measuring brain activity. EEG has several advantages over the other methods: its temporal resolution is higher and it directly measures the electrical activity of the brain. EEG has been a very useful clinical tool, especially in the field of epileptology, but also in other areas of neurology and psychiatry. EEG has been applied to monitor many clinical conditions such as seizure disorders, brain tumors, cerebrovascular disorders, metabolic and toxic encephalopathies, infections of the central nervous system, degenerative disease, head trauma, headache, brain death and electrocerebral inactivity [2].

EEG has received renewed research interest and has been applied to many neuronal pathological diseases because of its ability to capture the fluctuation of the electrical activity in the brain. The fluctuation may reveal important functional information about brain

processes. Higher brain processes depend upon the complex and most likely nonlinear interactions of multiple brain regions [3]. Even performing the most trivial tasks requires correlation of many widely distributed brain areas in the form of functional networks. EEG is one of the measurements that is sufficiently sensitive to capture this correlation in the brain. The ability to coordinate different parts of the brain has become increasingly recognized as it is important for normal brain functioning and may be impaired in some neurological diseases. A few studies have been done, using EEG, between healthy groups and disease groups with various neurological disorders such as Alzheimer's disease [4] and Schizophrenic disease [5]. They determined the differences between groups in the form of cortico-cortical connection and the transmission of information between different cortical areas.

The cortico-cortical connection in the brain is obtained through the application of statistical analysis . The analysis techniques require the signals to be stationary (ie. the amplitude and frequency content do not vary with time). However, EEG is a non-stationary signal. The non-stationary phenomena is caused by the switching of inherently metastable states of neural assembles during task performance that causes abrupt transitions [6]. The phenomenon is present in EEG in the form as alternation of relative homogenous intervals (segments) with different statistical features [7]. Although EEG is non-stationary over a long period of time, it may be considered locally stationary over short time intervals [8]. As a result, EEG data have to be divided into short time segments that are quasi-stationary (ie. statistically homogenous) in order to carry out further statistical analysis.

Recent research has explored statistical techniques that reveal the interdependencies between different parts of the brain in the form of functional connectivity network. In other words, the techniques attempt to determine whether there is any commonality between the time series of the EEG recordings as a sign of their relationship. If there indeed is an interdependency between two channels of EEG recordings, a connection is established. Based on the connections, functional connectivity network for the entire brain may be interpreted. Previous techniques that quantitatively relate the number of connections between neuronal

ensembles measure mainly the linear dependencies between EEG channels. The aim of this research is to apply nonlinear technique to obtain the functional connectivity networks from a normal group and a Parkinson's disease (PD) group and discriminate between them. PD is a form of neuronal degenerative disease that primarily affects the motor system. It is believed that PD is due to the impaired connectivity between different brain regions.

The rest of this chapter is organized as follows: section 1.2 presents an introduction to the EEG signals. This is followed by a literature survey on various EEG segmentation methods and EEG analysis techniques in section 1.3. Finally, the research objectives and methodology are presented in section 1.4.

1.2 Background on Electroencephalogram

EEG is generated by changes in the electrical charge of the nerve cells of the cerebral cortex. These neurons have a resting potential which is a difference in electrical potential between the interior of the cell and the extracellular space. The neurons are separated by a very small cleft. The cleft and the specialized membrane between two neurons is the synapse. The synapse is a polarized structure which allows the information to travel from one neuron to another. The resting potential fluctuates as a result of impulses or nerve stimulus arriving from other neurons at synapses. The impulses create the local postsynaptic potentials causing electrical current to flow along the membrane of the cell body. Once the changes reduce the membrane potential to a critical level, an action potential is generated. The action potential, which propagates along the neuronal membrane, is a transient disruption of the membrane potential. The synaptic activity gives rise to constantly changing ionic currents across the cell membrane and electrical potential field. Because of the briefness of the action potentials, the fluctuations in the surface EEG are produced mainly by the temporal and spatial summation of electrical currents caused by the relatively slow post-synaptic potentials. The electrical signals are manifestations of these complex systematic

physical and chemical processes which create these potential differences [1, 3, 9]

The majority of clinical EEG applications today use surface electrodes. The disks of the electrodes are fixed to the skin with electrode gel to form a malleable connection between the rigid disk and the skin. Electrodes are placed on the scalp according to the standard International 10-20 system. Most clinical applications use 19 recording electrodes, a ground and a system reference. Each electrode is connected to one input of a differential amplifier with the other input connected to the common system reference electrode. The reference electrode, usually the ear, is required to be relatively inactive compare to the cortical cortex activities. The purpose of the differential amplification is to eliminate the electrical potentials that do not come from the brain and are common to all electrodes. The amplified signals (typically 1,000∼100,000 times) are digitized by an analog-to-digital converter and typically filtered between 0.5∼1Hz and 35∼70Hz to remove the low frequency electrogalvanic and movement artifacts and high frequency electromyographic signals. A typical adult human EEG signal is about 10∼100μV in amplitude [1, 10].

The EEG signals are described in terms of rhythms and they are classified into frequency bands [10, 11]. A summary of EEG rhythms are provided in Table 1.1

Table 1.1: EEG rhythms

Rhythm	Frequency Rang	Description
Delta	<4Hz	Prominent feature of sleep in adult; The activity has a frontal predominance in adults, while in children the activity has an occipital predominance
Theta	4-7Hz	Drowsiness and sleep in adults and the waking state for young children
Alpha	8-13Hz	Normal, relaxed, awake with their eyes closed; the posterior dominant rhythm
Beta	>13Hz	Most evident in the frontal region; associated with active, busy or anxious thinking and active concentration

Artifacts are signals in EEG that are not caused by cerebral activity. EEG recordings often suffer from such artifacts as a result of physiological activities originating from ocular or body movements, muscle contractions, heartbeat, and external interference. Ocular artifacts are caused by the potential difference between the cornea and retina. The constant

and unavoidable reflexive eye movements generate a potential that is picked up in the frontal electrodes. Muscle activities caused by movements such as blinking or chewing may create a very short potential artifacts as well. EEG may also pick up artifacts from the heartbeat, especially in linkages across the head to the left ear; these artifacts may be mistaken for spike activity. In addition, external artifacts such as the ones from the power line or equipment might appear in EEG if they are not grounded properly. These artifacts usually have a frequency of 60Hz in North America [1].

1.3 Literature Survey

This section gives a review on the traditional EEG segmentation techniques and a number of bivariate methods frequently used for EEG analysis. Two of the commonly used EEG segmentation approaches are introduced in section 1.3.1. In section 1.3.2, the traditional linear methods used to assess interdependencies between EEG signals in the time and frequency domains are discussed. The nonlinear counterparts of the time domain methods are also discussed.

1.3.1 Traditional EEG Segmentation Techniques

Even though a long duration of the recorded EEG data is non-stationary, it may be considered quasi-stationary over a short interval. Researchers have explored various techniques to segment the EEG data. Segmentation of the EEG using fixed intervals and parametric segmentation are two common traditional EEG segmentation methods and they are briefly introduced.

Segmentation of the EEG using Fixed Intervals

The basic idea of this approach is to fit separate models to consecutive segments of a fixed and sufficiently small length. The approach assumes the duration of a minimal stationary

interval of the data. The EEG data are equally divided into the assumed fixed interval length. The segments are characterized by features such as spectral estimations or autoregressive coefficient [8]. Using a multivariate statistical procedure, the segments are classified into one of a number of classes according to their characteristics. Finally, the segments that belong to the same class are concatenated to form the stationary EEG signals where the parameters remain relatively constant over time.

However, the approach has some serious disadvantages. Because the duration of the stationary interval is assumed, the real boundaries of the stationary segments might lie inside the assumed interval. As a result, the EEG segments might contain transition processes and are not truly stationary. In order to overcome the boundary problem, an adaptive segmentation approach is developed.

Parametric Segmentation of the EEG

The general procedure of the adaptive parametric EEG segmentation is based on the estimation of the similarity of an initial fixed interval of EEG with an interval of the same duration viewed through the time window running along the entire EEG recording [8]. The similarity index would drop if the window ran over a transition between statistically inhomogeneous segments. The reduction of this similarity index is an indication of the segment boundary. The autoregressive model (AR) is one of the commonly used methods for this application. The AR model is a mathematical modeling of the EEG amplitude at a given moment based on the weighted sum of the EEG amplitude at prior moments. If $y(k)$ is the k-th value of the time series, the AR model of order N is given by:

$$\tilde{y}(k) = \sum_{i=1}^{N} a_i y(k-i) + n(k), \tag{1.1}$$

where $n(k)$ is the noise. Indeed if the stochastic nature of the time series follows certain dynamic rules and can be modeled, the AR model can predict the EEG amplitude with

certain accuracy. The differences between the predicted amplitude, $\tilde{y}(k)$, and the recorded EEG amplitude, $y(k)$, is the prediction error,

$$e_p = \tilde{y}(k) - y(k). \tag{1.2}$$

Within a stationary segment, the prediction error should be relatively constant and small. A decision about the boundary is apparent once a sharp increase in the prediction error is observed, indicating the termination of the current segment and the beginning of the next segment. A new set of model parameters can be derived from the initial portion of the next segment and the search for the next boundary can be continued.

1.3.2 Traditional Linear Analysis of the EEG

The bivariate techniques measure the pair-wise interdependencies between two EEG channels. Cross-correlation and coherence are two of the oldest and most classical analysis techniques in the time domain and the frequency domain, respectively. Granger causality is later developed for the study of the causal relation between EEG signals, which is not fully addressed by cross-correlation and coherence [12].

Cross-correlation

Cross-correlation measures the linear correlation between two time series (variables) X and Y as a function of their time delay (τ). The idea of time delay is significant because it may reflect a causal relationship between the signals. For example, if Y is caused by X, one may in principle derive a delay from X to Y. Nevertheless, the measured delay must be interpreted with caution because it might change due to the internal delay loops of one of the systems or different distances to the sources [13].

For signal $x(t)$ and $y(t)$ that are normalized with zero mean and unit variance, the

cross-correlation function is:

$$C_{xy}(\tau) = \frac{1}{N-\tau} \sum_{k=1}^{N-\tau} x(k+\tau)y(k), \tag{1.3}$$

where N is the total number of samples and τ is the time lag between the signals. C_{xy} ranges between -1 and 1. When $C_{xy} < 0$, it implies an inverse correlation and a tendency of both signals having similar absolute values but with opposite signs. When $C_{xy} > 0$, it implies a direct correlation and a tendency of both signals having similar values with the same sign. $C_{xy} = 0$ indicates the lack of the linear dependency between two signals. Usually, the τ that maximizes the cross-correlation function is used as an estimation of the delay between the signals. Cross-correlation at zero time lag is the well-known Pearson's product moment correlation coefficient and is also a popular measure for the EEG analysis.

Coherence

Coherence measures the linear correlation between two signals as a function of the frequency. It became a popular analysis method after the introduction of the fast Fourier transform (FTT) algorithm and has been applied to the study of EEG signals in several cognitive and clinical conditions. Coherence also known as the magnitude squared coherence is the cross-spectral density function between two signals. It is the Fourier transform of Eq. 1.3 normalized by auto-spectral density function of the individual signal. Due to the finite size of the EEG data, the spectra can only be estimated by averaging the periodogram over the epochs of the EEG signals (Welch's method). The coherence is

$$\kappa^2_{xy}(f) = \frac{|\langle S_{xy}(f)\rangle|^2}{|\langle S_{xx}(f)\rangle||\langle S_{yy}(f)\rangle|}, \tag{1.4}$$

where $\langle\ .\ \rangle$ indicates average over the number of epochs. The range of the estimated coherence is between 0 and 1. As coherence is a function of frequency, when $\kappa^2_{xy}(f) = 1$, it indicates that the signal are maximally dependent on each other at the particular frequency

f. When it is at 0, it means that the signals do not have any linear dependency at the particular frequency.

Granger Causality

One of the great interest in neurophysiology is to investigate the causal relation between two brain regions. While the time-delay between cross-correlation might provide some insight, inferring causality from the delay is not always straightforward [14]. As a consequence, Granger Causality was developed to address this particular issue. For two simultaneously recorded signal, if the prediction of the first signal is improved by including the past information of the second one, then the second signal is causal to the first one. Similarly, in the linear stochastic modeling of the times series framework [15], if X is causing Y, then by adding the past values of X to the regression of Y, the prediction of Y will be improved, which can be evaluated by comparing the univariate and bivariate fitting of the AR models to the signals. The univariate model has

$$\begin{aligned} x(n) &= \sum_{k=1}^{p} a_{xk} x(n-k) + u_x(n), \\ y(n) &= \sum_{k=1}^{p} a_{yk} y(n-k) + u_y(n), \end{aligned} \tag{1.5}$$

where a and b (in the bivariate case) are the model parameters, p is the model order, and u is the uncertainty or the residual noise associated with the model. The prediction error here depends only on the past values of itself. The bivariate model has

$$\begin{aligned} x(n) &= \sum_{k=1}^{p} a_{xyk} x(n-k) + \sum_{k=1}^{p} b_{xyk} y(n-k) + u_{xy}(n), \\ y(n) &= \sum_{k=1}^{p} a_{yxk} y(n-k) + \sum_{k=1}^{p} b_{yxk} x(n-k) + u_{yy}(n). \end{aligned} \tag{1.6}$$

Here, the prediction error depends on the past values of both signals.

For both models, the prediction performance can be assessed by the variances of the

prediction errors as,

$$V_{X|X_{-}} = var(u_x) \text{ and } V_{Y|Y_{-}} = var(u_y) \text{ for the univariate AR model}$$
$$V_{X|X_{-},Y_{-}} = var(u_{xy}) \text{ and } V_{Y|Y_{-},X_{-}} = var(u_{yx}) \text{ for the bivariate AR model,} \quad (1.7)$$

where $var(.)$ refers to the variance operator, $X|X_{-}$ and $X|X_{-}, Y_{-}$ indicate predicting X by its past values alone and by past values of both X and Y, respectively. If $V_{X|X_{-},Y_{-}} < V_{X|X_{-}}$, then Y causes X in the sense of Granger causality. The Granger causality defined for Y to X is

$$G_{Y \to X} = ln(\frac{V_{X|X_{-}}}{V_{X|X_{-},Y_{-}}}). \quad (1.8)$$

The Granger causality will be approximately zero if Y does not cause X (ie. $V_{X|X_{-},Y_{-}} \approx V_{X|X_{-}}$).

1.3.3 Nonlinear Analysis of the EEG

In the more recent years, the nonlinear techniques have been developed to exam the nonlinear interdependencies between EEG signals. Although nonlinear correlation coefficient was proposed about 15 years ago, the application has been mostly on the epileptic EEG data. Nonlinear Granger causality has been proposed as well, but this extension of the Granger causality has only been applied to simulated data and not yet to the real EEG signals.

Nonlinear Correlation Coefficient

Nonlinear correlation coefficient is a non-parametric nonlinear regression coefficient that measures the dependencies between two signals without making assumptions on the type of relationship between them [14]. The basic idea of this technique is that if the value of X is a function of the value of Y, the value of Y given X can be described by a nonlinear regression curve. The variance of Y according to the regression curve is the explained variance predicted by the knowledge of X. By subtracting the explained variance from the

original one, the unexplained variance can be estimated. The correlation ratio η^2 is the reduction of variance of Y obtained by predicting the Y values from those of X according to the regression curve:

$$\eta^2 = \frac{\text{(total vairance - unexplained variance)}}{\text{total variance}}. \tag{1.9}$$

The estimation of the above ratio is the nonlinear correlation coefficient h^2. In order to obtain this estimation, the values of X are subdivided into bins. For each bin, the value of the midpoint (p_i) and the average value of Y (q_i) are calculated. By connecting the points (p_i, q_i) by segments of straight lines, the regression curve is approximated. The nonlinear correlation coefficient between signals X and Y with the mean removed is thus:

$$h^2_{y|x} = \frac{\sum_{k=1}^{N} y(k)^2 - \sum_{k=1}^{N} (y(k) - f(x_i))^2}{\sum_{k=1}^{N} y(k)^2}, \tag{1.10}$$

where $f(x_i)$ is the linear piecewise approximation of the nonlinear regression curve. The nonlinear correlation coefficient ranges from 0 to 1, where zeros means complete independence and one means complete dependence. If $h^2_{y|x} = h^2_{x|y}$, it indicates that the relationship between these signals is linear.

Nonlinear Granger Causality

The original assumption of Granger causality is that the relationship between interacting systems is linear. Recently, researchers have attempted to modify the technique to incorporate the nonlinear properties of the signals [16]. However, there is no general framework of nonlinear models that can precisely model the dynamical characteristics of the neural signals. In [17], the author suggested an approximate solution to substitute the globally nonlinear model by a locally linear one. The nonlinear Granger causality has been proposed to be use for detecting the causal relation between two signals, but the effectiveness has not

yet been demonstrated in real data.

1.4 Research Objectives and Methodology

In this thesis, we examine the EEG signals collected during simple hand movements from normal subjects and PD subjects while they are off medication and on medication. The objectives of this research are:

1. to segment EEG data, so it is task-related and quasi-stationary,
2. to apply nonlinear analysis on the segmented EEG data and derive the functional connectivity networks between different subject groups, and
3. to compare the network differences between normal subjects and PD subjects (off and on medication).

To achieve these objectives, we propose a novel segmentation method based on the temporal dynamics of the cross-spectrogram of the computed Independent Components (ICs) of the EEG signals. We apply mutual information (MI) as an alternative metric of assessing possible nonlinear statistical dependencies between the time-series of the EEG signals. In a real case study of PD, graphical theoretical analysis is applied to the MI based networks and one-way analysis of variance (ANOVA) test is applied to the group analysis with a task factor and a subject group factor.

The rest of the thesis is organized as follows: chapter 2 presents the detail of a novel segmentation, mutual information network analysis and mainly focuses on the comparison between normal subjects and PD subjects - off medication. In chapter 3, the additional results including comparison between normal subjects, PD subjects on and off medication are given. Lastly, we conclude the work along with suggestions for future work in chapter 4. The main focus of this thesis in on the mutual information network analysis of the EEG. As another project, spectral clustering of fMRI Data within regions of interest (ROI) has

been studied. The aim for this project is to identify the active regions of the brain that are task-related. We propose a spectral clustering method for assessing those voxels within an ROI that are suitable for further task-activation analysis. Similar to the EEG analysis, a real fMRI case study in Parkinson's disease is investigated and the comparisons between different subject groups are made. We have included spectral clustering of the fMRI data as appendix A at the end.

Chapter 2

A Novel Segmentation, Mutual Information Network Analysis [1]

2.1 Introduction

Connectivity between brain regions is being increasingly recognized as important for normal brain functioning, and may be impaired in some neurological diseases such as Parkinson's disease (PD). The electroencephalogram (EEG), with its excellent temporal resolution (∼1msec), is the most widely available technology used for inferring transient synchronization in brain studies. Both linear and nonlinear measures have been applied to assess the interdependencies between EEG channels [12]. For example, coherence and correlation methods, which measure the dependencies between two EEG signals in the frequency and time domains respectively, have been applied to the EEG to study the cortical synchrony that can be modulated as a function of task, and may systematically differ between normal and disease groups [18, 19]. Nevertheless, they consider only linear dependencies and may be particularly sensitive to outliers.

An alternative to the linear methods of coherence and correlation is to consider the

[1] Versions of this chapter have been submitted for publication.

1. ©[2007] IEEE. Reprinted, with permission, from Pamela W. H. Lee, Z. Jane Wang, and Martin J. McKeown, "Mutual Information based Relevance Network Analysis: A Parkinson's Disease Study," in *Proc. 32th IEEE Int. Conf. Acoustics, Speech, and Signal Proc. (ICASSP32), 2008*, submitted for publication.

2. ©[2007] IEEE. Reprinted, with permission, from Pamela W. H. Lee, Z. Jane Wang, and Martin J. McKeown, "A Novel Segmentation, Mutual Information Network Framework for EEG Analysis of Motor Tasks," *IEEE Transactions on Biomedical Engineering*, submitted for publication.

mutual information (MI) between channels over a given window. MI measures both the linear and nonlinear statistical dependencies between time series and thus can be used to detect the functional coupling. MI measures the amount of information x, generated from signal X, provides about y, generated from signal Y. It thus quantifies the information transmitted from one time series to another. It has maximum value when the two time series are the same and if the two time series are completely independent, the value is zero. Previously, researchers have applied MI on various pathological signals to gain a better understanding of the functional connectivity network. In [4], by estimating the MI between the time series of multiple pairs of EEG channels, they investigated the information transmission between different cortical areas in Alzheimer's patients. Similar studies have used MI as a marker for cortico-cortical connections in schizophrenic patients [5] and odor stimulation [20].

Prior studies investigating differences in the MI between EEG channels in patient populations have only used short segments ($\sim$16s) so that the assumption of stationarity is approximately valid. Also, previous studies have just had subjects lie passively with their eyes closed. While this may be adequate for subjects with diseases that directly affect the cortex from which the EEG signal is derived (e.g. Alzheimer's disease), measurements recorded from passive subjects will be less sensitive for neurodegenerative diseases that affect mobility, such as PD. PD is a form of movement disorder and is characterized by muscle rigidity, tremor, and bradykinesia (slowing of physical movement). These symptoms likely do not reflect a primary failure of the cortex, but rather the effects of failure of the basal ganglia to prime the cortex for preparation and execution of movement. As a result, PD patients have a difficult time performing simultaneous movements comparing to normal subjects [21, 22].

In order to assess the indirect effects on the cortex of basal ganglia dysfunction in PD, it is necessary to have subjects perform a task, such as a motor task. However, as soon as a subject performs a motor task, the non-stationarity nature of the EEG must be

considered [6]. The non-stationarity likely reflects the switching of inherently metastable states of neural assemblies during task performance causing abrupt transitions. In order to obtain the quasi-stationary segments in EEG signals and the synchrony between quasi-stationary structures in pairs of EEG channels, we propose a novel segmentation method of the EEG based on the temporal dynamics of the cross-spectrogram of the Independent Components (ICs). We apply graph theoretical analysis [23, 24] to the network of each group by computing the clustering coefficient (C) and shortest path length (L), similar to [25]. The network differences between tasks and between subject groups are also analyzed by one-way analysis of variance (ANOVA). To our knowledge, this is the first application of using transient synchronization of ICs for temporal segmentation of time series. Also, it is the first application of joint MI and network analysis to assess information transmission between different cortical areas in PD.

The chapter is organized as follows: the detail discussion of the methods is presented in section 2.2. Section 2.3 introduces the EEG experiment design and summarizes the results in a real case study of PD. Section 2.4 summarizes and concludes the chapter.

2.2 Methods

We recorded five minutes of data with subjects performing simple hand movements in order to gain insight into the difficulty that PD subjects face when performing simultaneous movements.

2.2.1 Preprocessing

Fig. 2.1 presents flowchart diagram of the steps of EEG preprocessing. The preprocessing includes bandpass filtering, Independent Component Analysis (ICA) noise removal, and EEG segmentation based on the cross-spectrogram of the ICs.

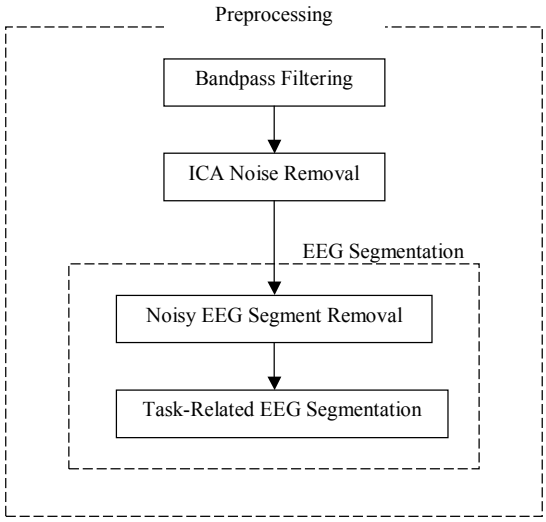


Figure 2.1: Flowchart of the Steps of EEG Preprocessing.

Bandpass Filtering

EEG data contain a wide range of frequency components, yet not all of them are of clinical and physiological interest. The data are therefore bandpass-filtered by a 4th order Butterworth filter between 0.5∼55Hz in order to focus on the frequency range of interest [26].

ICA Noise Removal

ICA has been proven capable of isolating both artifactual and neurally generated EEG sources [27]. As various contaminants of EEG recordings such as eye movements, eye blinks, cardiac signals, muscle contamination, etc., can be considered temporally independent from ongoing brain activity, ICA is useful for artifact removal. ICA decomposes mixture of time courses into a sum of temporally statistical maximally independent components. The EEG measurements from the scalp, $x = x_1(t), ..., x_N(t)$, are mixtures of the source signals, $s = s_1(t), ..., s_N(t)$, as

$$x = As, \tag{2.1}$$

where A is the unknown mixture matrix. The task of ICA is to recover a version, u, of the original sources, s, by finding an unmixing matrix, W, specifying spatial filters that invert the mixing process linearly, as

$$u = Wx. \tag{2.2}$$

Here the infomax-ICA [28] is applied to decompose the EEG signals. ICA finds a coordinate frame in which the data projection has minimal temporal overlap by minimizing the mutual information among the data projections or maximizing the joint entropy of a nonlinear function of s. It is most appropriate to perform ICA decomposition on sources that are linearly mixed in the recorded signals without time delays. ICA is applied to the EEG data to obtain the maximally independent time courses of different brain and artifact sources and the signals are "corrected" by eliminating the artifactual sources. As an example, Fig. 2.2 displays the ICs. Clearly, component 1 is an artifact unrelated to brain activity. This is confirmed by looking at the spatial weighting of the component. The component has maximum weighting at EEG channel FP1 and FP2 which also indicates to us that it is in fact an artifact related to the eye blinks and thus is removed.

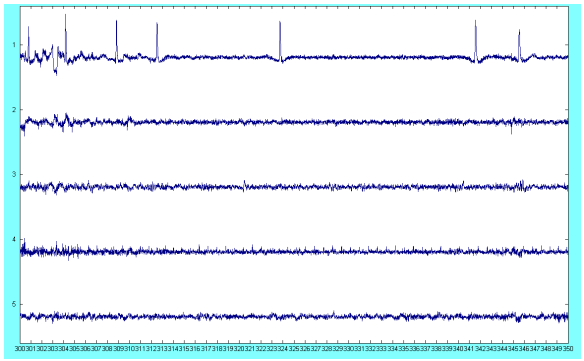


Figure 2.2: An example of the resulting ICs from the ICA decomposition of the EEG data. Component 1 is identified as a noise artifact.

EEG Segmentation based on Cross-Spectrogram

Due to the inherent non-stationary nature of EEG data [6], EEG segmentation based on the temporal dynamics of the cross-spectrogram of the ICs can be applied. If the derived com-

ponents were truly independent, then the cross-spectrum would not be significant. However, in real data many of the assumptions of ICA are violated. The data are not stationary, and the time courses are not temporally white. By using infomax-ICA, which does not incorporate time delays, the derived components will be maximally independent at zero lag. As such, it will deal with the problem of volume conduction – where a deep electrical source may impart common electrical activity to two or more electrodes. Even though ICs are maximally independent over the whole time range, they may exhibit partial synchronization within specific time/frequency window [29], through which the transient coupling of neural networks might be revealed. By examining the ICs within a short moving window, the non-stationary nature of the EEG will be explored, and significant dependencies between ICs become apparent. Recent studies such as [30–32] have also explored the transient synchrony between ICs and suggested transient correlation between ICs.

The ICs are thus transformed into time-frequency domain and the cross-spectrogram of every pair of ICs is computed. The frequency contents are computed by cross power spectral density using the Welch's averaged, modified periodogram [33] method of spectral estimation. If x(t) and y(t) are signals normalized to zero mean and unit variance, their cross-correlation is:

$$R_{xy}(\tau) = \frac{1}{N-\tau} \sum_{k=1}^{N-\tau} x(k+\tau)\, y(k)\,. \tag{2.3}$$

The cross spectral density function is a Fourier transform of the cross-correlation function that indicates the relationship between the two signals,

$$S(f) = \int_{-\infty}^{\infty} R_{xy}(t) e^{-j2\pi ft} dt. \tag{2.4}$$

By using a short-term time shifting window, we are able to obtain localized frequency content of the two signals and their relationship with respect to time and frequency. The cross-spectrum is computed based on a short (3s) time window shifted by 0.5s to obtain the localized time information. Based on the cross-spectrogram, we can detect and remove the

transient broadband artifacts by looking at the frequency band that does not contain any clinical or physiological interest. On the contrary, by examining the cross-spectrogram that explores the correlation between ICs at the frequency band of interest, we can identify the task-related segments.

2.2.2 Mutual Information based Relevance Network

MI measures the mutual dependence or information gained about one signal from another. The detailed derivation and background of information theory can be found in [34]. Here we give a brief derivation of the MI. Given a random variable X, the entropy H is the average amount of information obtained from any observation of X. This information is called uncertainty before any measurement of X,

$$H(X) = -\sum_{x} P_X(x) log_2 P_X(x), \tag{2.5}$$

where $P_X(x)$ is the probability that x is drawn from X and the probability distribution of values observed for the measurement x. If there is another signal Y, then the conditional uncertainty on X is introduced. Under the condition $Y = y$, $H(X)$ becomes the conditional uncertainty on X,

$$H(X|Y = y) = -\sum_{x} \frac{P_{XY}(x, y)}{P_Y(y)} log_2 \frac{P_{XY}(x, y)}{P_Y(y)}, \tag{2.6}$$

where $P_{XY}(x, y)$ is the joint probability density for the measurements of X and Y that produce the values x and y. The mean conditional uncertainty on X over y, under the condition that Y is known is thus

$$H(X|Y) = \sum_{y} P_Y(y) H(X|Y = y) \tag{2.7}$$

$$= -\sum_{x,y} P_{XY}(x,y) log_2 \frac{P_{XY}(x,y)}{P_Y(y)} \tag{2.8}$$

$$= H(X,Y) - H(Y), \tag{2.9}$$

where

$$H(X,Y) = -\sum_{x,y} P_{XY}(x,y) log_2 P_{XY}(x,y). \tag{2.10}$$

The conditional entropy measures the amount of uncertainty in X given that signal Y is known. $H(X,Y)$ is the joint entropy that is the joint uncertainty in X and Y. Generally, the conditional entropy will be less or equal to the individual entropy, because knowing another signal Y can only reduce the uncertainty in X. It is equal if and only if the two signals are independent. With the priori uncertainty on X, $H(X)$ and the posteriori uncertainty on X given Y, $H(X|Y)$, the amount of uncertainty in X that is reduced by knowing Y is the MI,

$$I(X,Y) = H(X) - H(X|Y) \tag{2.11}$$

$$= H(X) + H(Y) - H(X,Y), \tag{2.12}$$

which can also be written as

$$I(X,Y) = \sum_{x,y} P_{XY}(x,y) log_2 \frac{P_{XY}(x,y)}{P_X(x) P_Y(y)}. \tag{2.13}$$

MI is also the amount of information about X that Y contains. It is a symmetric function meaning $I(X,Y) = I(Y,X)$. A MI at zero means that Y does not contain additional information about X that is $H(X) = H(X|Y)$ or $H(X,Y) = H(X) + H(Y)$. If X is independent of Y, then $P_{XY}(x,y)$ factorizes to $P_X(x)P_Y(y)$ and the MI is zero. On the contrary, the higher the MI between two signals, the more information they contain about

each other. Hence, the higher MI, the more likely that the two signals are biologically related. MI is estimated from a finite number of samples, the probability densities, $P_X(x)$ and $P_{XY}(x, y)$ are approximated by histogram.

Relevance networks, originally devised for graphically depicting the relationship between genes [35] can be generalized to take large data sets of experimental data and graphically depict the result of pair-wise MI. Here we use relevance networks to describe the MI between each pair of EEG electrodes.

2.2.3 Network Analysis

Graph theoretical analysis is applied to the MI matrices of all possible pair-wise combinations of electrodes. The resulting MI matrices are converted to binary graphs by applying a threshold. A graph, which consists of nodes (vertices) and connections (edges), is a basic representation of a network. If two vertices are connected, then there exists an edge between them. If the vertices are connected one another by an edge, then they are "neighbors". The number of edges that have to be traveled from one vertex to another is defined as the distance between two vertices. Graphs are characterized by a cluster coefficient C and a characteristic path length L. The C of a graph is the ratio of all existing edges between the neighbors and the maximum possible number of edges between the neighbors. It ranges from 0 to 1 and measures the local connectivity of the graph. The L of a graph is the mean of all shortest paths (shortest distance) connecting all pairs of vertices. It has a value greater than 1 and measures the global connectivity of the graph. A detail graphical explanation of a graph and graph theoretical measures can be found in [25].

The full 19x19 MI matrices are converted to unidirectional binary graph by considering a threshold T. The edges between two vertices do not have graded values; in another words, they either exist or do not exist. We have computed the C and L as a function of T which ranges from 0.01 to 0.3 with increment of 0.002. In the analysis, we compared the overall C which is the average of the C over all the vertices within the graph with L and also look

at the C per vertex in each group. However, when C and L are computed as a function of T, the results might be biased by the mean level of MI between each group. To control this effect, we turn to another measure called graph degree K which is the average number of edges per vertex. Since the graphs have the same number of edges with the same K, the remaining differences would reflect the actual differences in graph organization.

The MI differences between tasks and between subject groups are also analyzed by ANOVA [36]. The normality of the distribution of the MI values is tested by the Kolmogorov-Smirnov (KS) test. Within group differences of each MI value are analyzed using ANOVA with a task factor. Across group differences of each MI value are analyzed using ANOVA with a subject group factor. Because we are interested in the significant connections with greater magnitude, a threshold is further applied to filter the connections. In order to automatically select the threshold value, we adapt a statistical test based on permutation [37], which places no a priori hypothesis on the signals. The permutation results are generated by randomly permuting the order of the second signal and computing the pair-wise MI based on them. The connections between any two electrodes are established for the relevance network if they are significantly different according to the ANOVA test and have magnitudes that are greater than the permutation threshold.

2.3 Experiment and Results

2.3.1 Subjects and Experiment Design

All research was approved by the University of British Columbia Ethics Board. After giving informed consent, seven PD and six age-matched control subjects volunteered to participate in the study. All patients were diagnosed with mild to moderately severe PD (Hoehn and Yahr stage 1-3) [38]. The control subjects were confirmed to be without active neurological disorders. All patients were taken off L-Dopa medication after overnight withdrawal of at least 12 hours. The experiment was performed in the off medication state and then repeated

1 hour after immediate-release formulation of L-Dopa.

Subjects were asked to hold a custom-built rubber squeeze bulb in their right hand with their arm stabilized in one position. Each subjects had their maximum voluntary contraction (MVC) tested at the start of the experiment and all subsequent forces were scaled to this accordingly. They were instructed to control an "inflatable" ring as shown as the horizontal bar in Fig. 2.3 by squeezing the bulb. The ring must move through an undulating tunnel without touching the sides. The required pressure was between 5∼15% MVC in order to successfully avoid the sides of the tunnel. Two five-minute trials were performed by normal subjects (N), PD before medication (Ppre), and PD 1hr after L-Dopa mediation (Ppost). During one trial, subjects were asked to squeeze the bulb (SQ) with right hand alone. In another trial, subjects were asked to press a mouse button intermittently with their left hand while they observed a color changes in the ring in addition to doing the squeezing of the right hand, exactly as before. We were particularly interested in how subjects fared when they were required to do both movements simultaneously (BO), as seen in Fig. 2.3, compared to the squeezing alone (SQ) condition.

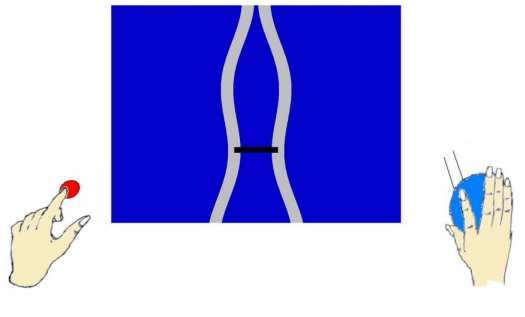


Figure 2.3: Experiment design of SQ and BO task. In the middle is the "inflatable" ring indicated by the horizontal bar. Two separate tasks are performed: 1) control the "inflatable" ring by squeezing the bulb with right hand alone; 2) press a mouse button intermittently with their left hand in addition to doing the squeezing with the right hand.

2.3.2 EEG Data Preprocessing

Subjects wore an electrode cap (Electro-Cap International, Eaton, OH) and the data were collected using a Ceegraph Netlink system from Bio-Logic Systems (Illinois). The 19 EEG electrodes/channels are shown in Fig. 2.4 and the data are sampled at 128Hz and bandpass-filtered between 0.5∼55Hz. The electrodes correspond to the channel number from 1 to 19 starting from the left to right and top to bottom. The artifactual components were screened and removed from the data by visual inspection. Two operations are done based on the cross-spectrogram of the ICA components: noisy EEG segment removal and task-related EEG segmentation.

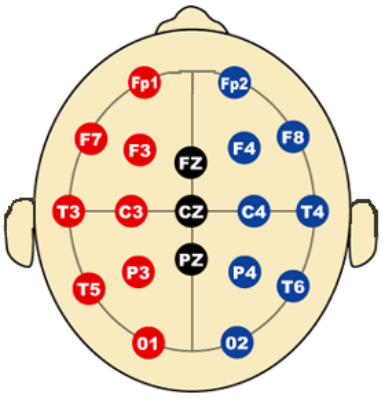
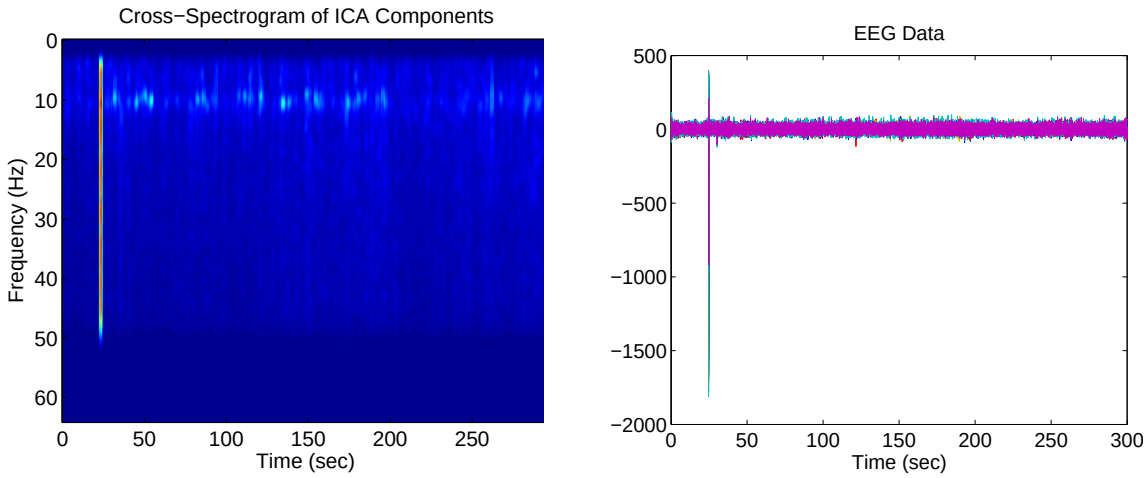


Figure 2.4: EEG 10-20 Electrode Placement. The electrodes correspond to the channel number from 1 to 19 starting from the left to right and top to bottom. Abbreviations: F = frontal, C = central, P = parietal, T = temporal, O = occipital, Fp = frontopolar.

Noisy EEG Segment Removal

The data in the frequency range of 45∼55Hz do not typically contain information of clinical or physiological interest. Thus, activity in the cross-spectrogram of the ICs between 45∼55Hz is a good marker of transient broadband artifacts that were not eliminated by ICA Noise Removal step. As can be seen in Fig. 2.5, the broadband artifact right around 23s of the cross-spectrogram of the ICs reflects the temporal noisy segment in the actual EEG data. Thresholding is applied to eliminate EEG segments with greater power in this frequency range and exclude them from further analysis.



(a) Cross-Spectrogram of a pair of ICs (b) Corresponding EEG Data

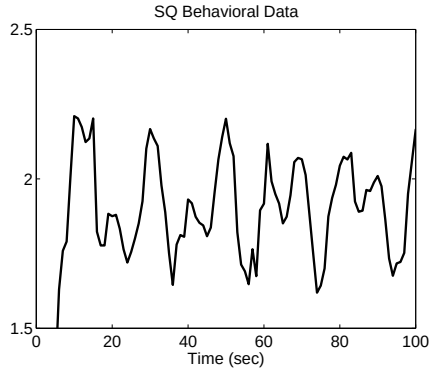
Figure 2.5: An example of the cross-spectrogram and EEG. A broadband artifact (at around 23s) exists in the cross-spectrogram of the ICs which reflects the temporal noisy segment in the actual EEG data.

Task-Related EEG Segmentation

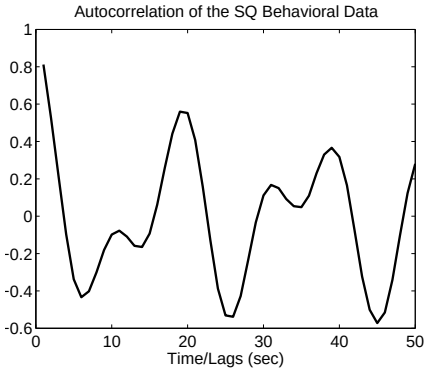
The segmentation and network analysis are based on three frequency bands: 5∼8Hz (Theta band), 8∼12Hz (Alpha band), 12∼30Hz (Beta band) as these have in the past been implicated with different neurological processes [26]. Since not all IC pairs are task-related, only the ones that are modulated with the behavioral data are used. From the behavioral data we see that SQ contains two sinusoids with period of 10 and 18 seconds. For the button press, color changes occur every 20 seconds. Therefore, the autocorrelation of cross-spectrogram of the ICs over these three frequency bands that has a peak at 10 seconds or 18∼20 seconds or both is selected for the EEG segmentation of SQ and BO.

Because these IC pairs that are modulated with the behavioral data have a common task-related feature, they are integrated to enhance the signal-to-noise for segmentation. As an example, both the autocorrelation of the behavioral data and the autocorrelation of the integrated cross-spectrogram IC pairs contain peaks around every 10 and 18∼20 seconds as seen in Fig. 2.6(b) and 2.6(d), yet there are some discrepancies between their actual time courses in Fig. 2.6(a) and 2.6(c). In general, the behavioral data cannot

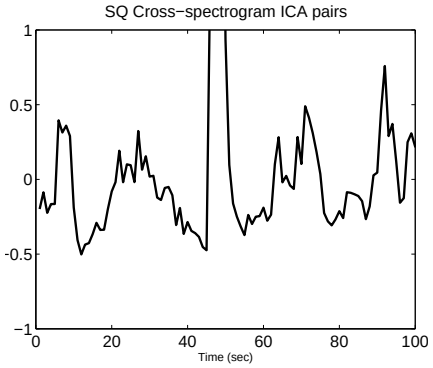
reflect exactly how the brain functionally relates to the task, although in this case, we can use the smooth behavioral performance as a rough check on the validity of the proposed segmentation approach. EEG segmentation is thus based on the actual IC pairs directly for more accurate segmentation. Depending on the features of each dataset, approximately five pairs are chosen for each task. The EEG segments that have higher power indicate that they are highly correlated for the particular task they are performing. We therefore threshold the EEG data every 20 seconds to obtain the task-related segments. Only segments that are above the mean plus the mean absolute deviation are selected for further analysis.



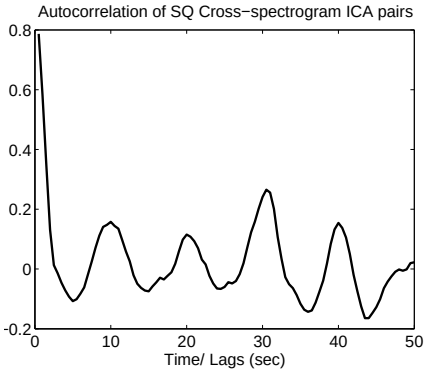
(a) SQ task: the behavioral data



(b) SQ task: the autocorrelation of the behavioral data



(c) SQ task: the sum of all behaviorally relevant cross-spectra between IC pairs

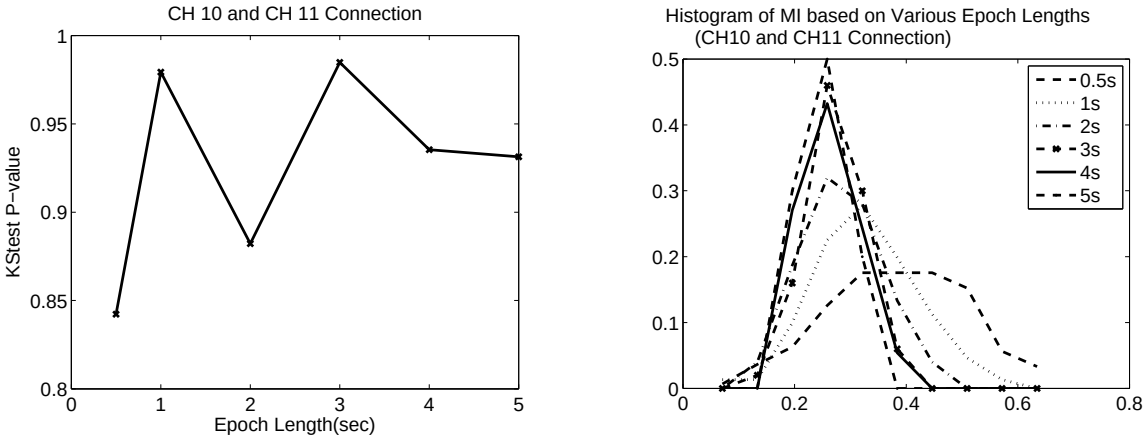


(d) SQ task: the autocorrelation of the sum of cross-spectra between IC pairs

Figure 2.6: An example of the original and autocorrelation of the SQ behavioral data and the sum of cross-spectrogram IC pairs integrated between 8∼12Hz. Both the autocorrelation of the behavioral data and the autocorrelation of the integrated cross-spectrogram IC pairs contain peaks around every 10 and 18 seconds, yet there are some discrepancies between their actual time courses.

2.3.3 Mutual Information Analysis

The data are separated into 4 second epochs for MI computation in order to increase the sample size as well as to enhance the stationarity and consistency of the MI estimates. In our primary study, we have tried varying the length of the epochs and observed that 4 second epochs give a more Gaussian distribution. The p-value from KS-test versus various lengths of one connection is shown in Fig. 2.7(a) as an example. It can be seen that at around 1∼4 second the data become more Gaussian. We have chosen 4 second as it is both Gaussian and gives us enough data points for more accurate MI computation. The distribution of MI values of the connection based on various lengths is also illustrated in Fig. 2.7(b).



(a) Normality (p-value) of MI based on different epoch lengths

(b) Distribution of MI based on different epoch lengths

Figure 2.7: Relationship between the epoch length and normality. The data are more Gaussian for epoch length between 1 to 4 second observed both from the p-value of the KS-test and the distribution of the data.

Graphical theoretical measure is applied to each network. The MI matrix for each subject is converted to a graph separately, and the means of the overall C and L within the group (ie. N-SQ, Ppre-SQ, N-BO , and Ppre-BO) are computed as a function of T. In order to check the mean level differences between each group, the relationship between graph degree K and T are investigated. We find that K as a function of T is approximately the same for all cases, and one illustration is shown in Fig. 2.8. This suggests to us that the variances

of the mean level for each group are very small. To keep our analysis simple, our analyses are done as a function of T.

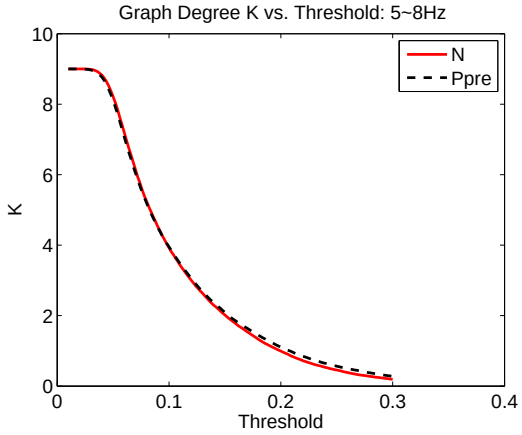


Figure 2.8: An illustration of the graph degree K as a function of threshold T. K as a function of T is approximately the same for all cases.

As T is varied from 0.01 to 0.3, the graphs start to break into subgraphs. In addition, at higher T, some subjects start to have empty graphs meaning the graphs contain no connection at all. Therefore, when we interpret the results, we need to make note of where those points are and they are summarized in Table 2.1 and 2.2. Again, we see that the points between N and Ppre do not differ much because their means are very close.

Table 2.1: SQ Task: Threshold T when graphs start to split into subgraphs or become empty graphs

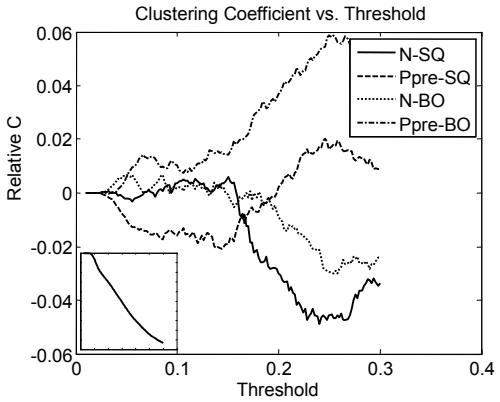
	N-SQ		Ppre-SQ	
	Subgraphs	Empty Graphs	Subgraphs	Empty Graphs
5∼8Hz	0.0600	0.2340	0.0660	0.2380
8∼12Hz	0.0520	0.2140	0.0560	0.2120
12∼30Hz	0.0500	0.2060	0.0520	0.1660

Table 2.2: BO Task: Threshold T when graphs start to split into subgraphs or become empty graphs

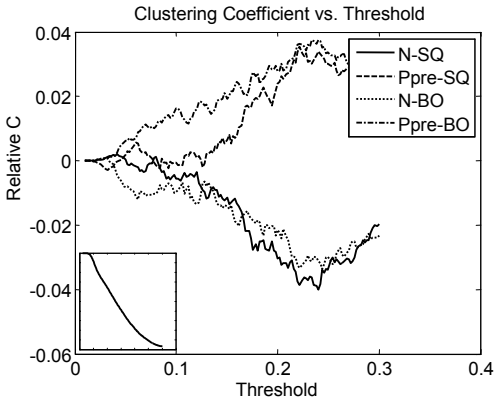
	N-BO		Ppre-BO	
	Subgraphs	Empty Graphs	Subgraphs	Empty Graphs
5∼8Hz	0.0760	0.2600	0.0600	0.2360
8∼12Hz	0.0460	0.1980	0.0580	0.1980
12∼30Hz	0.0440	0.1920	0.0540	0.1640

The mean of the overall C for each group as a function of T is computed and compared.

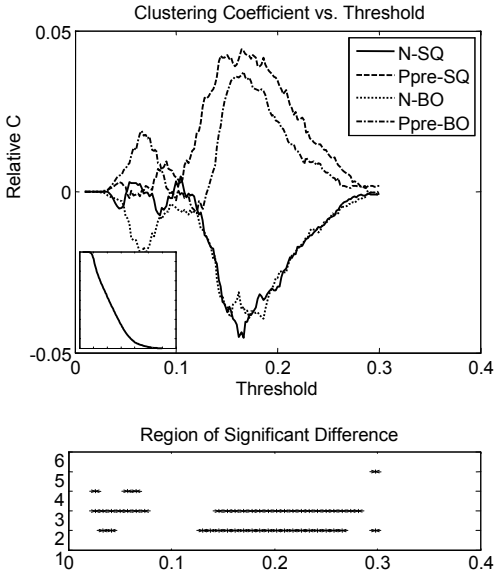
Because the means of the four groups (ie. N-SQ, Ppre-SQ, N-BO, Ppre-BO) as a function of T follow the same pattern and we are more interested in the differences between groups, the deviation from the overall mean of the four groups as a function of T is illustrated at the top panel of Fig. 2.9. The overall mean of the four groups as a function of T is displayed at the bottom left corner of the top panel. The bottom panel shows the region that is significantly different between groups (at 2: N vs. Ppre (SQ); at 3: N vs. Ppre (BO); at 4: SQ vs. BO (N); at 5: SQ vs. BO (Ppre)). The significance of the overall C between groups is tested by the pair-wise T-test ($p < 0.05$). In general, the intensity of the cluster does not differ significantly between tasks (SQ and BO) within a group, except for Ppre at lower frequency band (5∼8Hz). The graphs are more clustered for Ppre compared to N for SQ task and BO task, especially at higher T ($T > 0.15$). For Ppre subjects, while they are performing BO task, the graphs are intensely clustered in the lower frequency band (5∼8Hz) and the medium frequency band (8∼12Hz). The consistently lower C seen in N across all frequency bands probably reflects the widespread, excessive synchronization seen in PD [39–41]. Unlike prior studies that have emphasized synchronization in the beta band, we have observed excessive synchronization in all bands.



(a) 5∼8 Hz



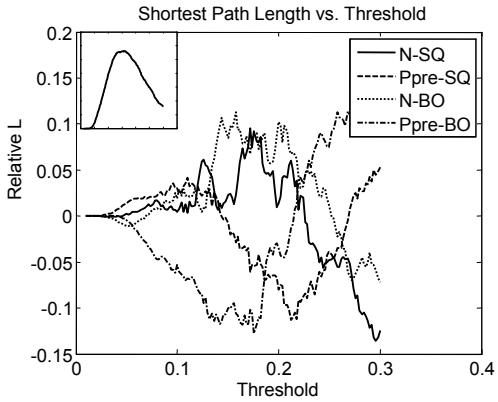
(b) 8∼12 Hz



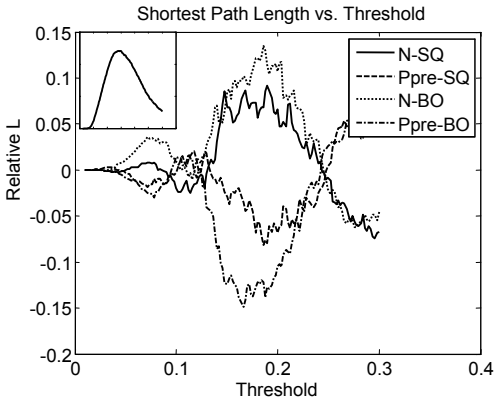
(c) 12∼30 Hz

Figure 2.9: The deviation of C from the overall mean of the four groups as a function T. The overall mean of the four groups is presented at the bottom left corner of the top panel. The bottom panel indicates the region that is significantly different between groups (at 2: N vs. Ppre (SQ); at 3: N vs. Ppre (BO); at 4: SQ vs. BO (N); at 5: SQ vs. BO (Ppre)).

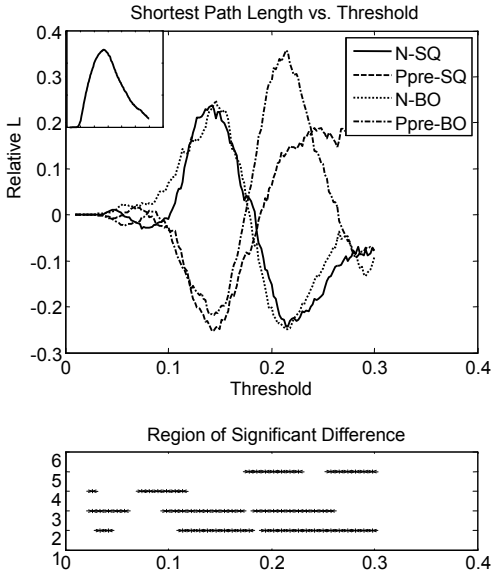
The top panel of Fig. 2.10 shows the deviation of L from the overall mean of the four groups as a function T. The overall mean of the four groups is presented at the top left corner of the top panel. The region that is significantly different between groups (at 2: N vs. Ppre (SQ); at 3: N vs. Ppre (BO); at 4: SQ vs. BO (N); at 5: SQ vs. BO (Ppre)) is shown at the bottom panel. Here, the L does not have much difference for N performing either SQ task or BO task. The L of N is quite a bit larger than the one of Ppre between $\sim$ 0.1 to 0.25, but L of Ppre becomes quite a bit larger than the one of N after that. However, the actual implication of the L must be taken with caution as the region of significance lies in the subgraphs region. Overall, Ppre subjects with larger C and shorter L compared to N subjects might suggest that the graphs are more broken up into small tightly connected clusters.



(a) 5∼8 Hz



(b) 8∼12 Hz



(c) 12∼30 Hz

Figure 2.10: The deviation of L from the overall mean of the four groups as a function T. The overall mean of the four groups is presented at the top left corner of the top panel. The bottom panel indicates the region that is significantly different between groups (at 2: N vs. Ppre (SQ); at 3: N vs. Ppre (BO); at 4: SQ vs. BO (N); at 5: SQ vs. BO (Ppre)).

In order to see where the clusters are located, the C for each vertex/channel is shown in Table 2.3 and 2.4 and the significance is checked by pair-wise T-test. For SQ, clusters are intensely formed on the left side of the frontal and motor cortex for Ppre over all frequency bands. Whereas, for N, they are more intensely formed on the left side of the posterior cortex. The strong coupling in the left hemisphere is consistent with the SQ task being performed by the right hand. Similar results are observed for BO. The resulting networks involve both hemispheres of the brain since BO is performed by both the left and right hand.

Table 2.3: SQ Task: C per vertex and the p-value of the pair-wise T-test between two groups. (T = 0.2 for 5∼8Hz; T = 0.18 for 8∼12Hz; T = 0.15 for 12∼30Hz)

	5∼8Hz			8∼12Hz			12∼30Hz		
Channel #	N mean	Ppre mean	P-value	N mean	Ppre mean	P-value	N mean	Ppre mean	P-value
1	0.6471	0.6158	3.60E-01	0.6788	0.6752	8.98E-01	0.7395	0.7204	3.54E-01
2	0.5786	0.6075	4.42E-01	0.6229	0.7133	4.97E-03	0.6982	0.7660	6.75E-03
3	0.2106	0.2770	9.33E-02	0.2705	0.3638	2.58E-02	0.3961	0.5643	5.95E-05
4	0.5290	0.6062	3.28E-02	0.6170	0.6470	3.35E-01	0.6781	0.7039	2.37E-01
5	0.4791	0.6041	6.92E-06	0.5648	0.6767	6.15E-06	0.6439	0.7440	2.82E-07
6	0.5204	0.6574	8.34E-05	0.5363	0.6728	4.11E-06	0.6148	0.7013	1.39E-04
7	0.1958	0.2478	1.40E-01	0.2806	0.3679	2.95E-02	0.4731	0.4819	8.31E-01
8	0.0159	0.0630	7.14E-03	0.0312	0.1073	1.20E-03	0.1767	0.2333	1.10E-01
9	0.2138	0.2895	2.82E-02	0.3395	0.3392	9.92E-01	0.5177	0.4640	7.19E-02
10	0.2229	0.3275	9.49E-04	0.3466	0.3913	1.35E-01	0.5209	0.5234	9.20E-01
11	0.2530	0.3023	1.76E-01	0.3593	0.4065	1.87E-01	0.5156	0.5022	6.29E-01
12	0.0054	0.0739	2.38E-04	0.0453	0.1346	8.02E-04	0.1734	0.2746	4.57E-03
13	0.1844	0.1034	1.06E-02	0.2824	0.1384	7.93E-05	0.4799	0.2557	8.59E-08
14	0.2241	0.1917	3.05E-01	0.3117	0.2322	1.17E-02	0.4765	0.3414	7.11E-06
15	0.2409	0.1787	3.92E-02	0.3340	0.2310	8.52E-04	0.5111	0.3665	8.04E-07
16	0.2139	0.1838	3.25E-01	0.2691	0.2423	3.85E-01	0.4560	0.3563	6.30E-04
17	0.1123	0.1057	8.14E-01	0.2228	0.1265	3.59E-03	0.4521	0.2778	2.94E-05
18	0.1310	0.1537	4.61E-01	0.2029	0.1955	8.25E-01	0.4547	0.3436	4.26E-03
19	0.1428	0.1402	9.35E-01	0.2083	0.1900	5.93E-01	0.4615	0.3422	1.86E-03

Table 2.4: BO Task: C per vertex and the p-value of the pair-wise T-test between two groups. (T = 0.2 for 5∼8Hz; T = 0.18 for 8∼12Hz; T = 0.15 for 12∼30Hz)

	5∼8Hz			8∼12Hz			12∼30Hz		
Channel #	N mean	Ppre mean	P-value	N mean	Ppre mean	P-value	N mean	Ppre mean	P-value
1	0.6535	0.6496	9.04E-01	0.7140	0.7028	6.74E-01	0.7743	0.7636	5.61E-01
2	0.5988	0.6179	5.83E-01	0.6801	0.7089	3.26E-01	0.7158	0.7630	2.37E-02
3	0.1758	0.3215	8.66E-05	0.2286	0.3980	1.72E-05	0.3552	0.5811	1.91E-08
4	0.5665	0.6533	8.77E-03	0.6233	0.6993	9.53E-03	0.6844	0.7434	5.65E-03
5	0.5188	0.6053	1.40E-03	0.5661	0.6955	1.14E-08	0.6334	0.7611	4.61E-12
6	0.4917	0.6331	1.03E-05	0.5972	0.6808	1.75E-03	0.6668	0.7204	7.50E-03
7	0.2554	0.3190	9.12E-02	0.3399	0.4300	2.19E-02	0.5210	0.6103	1.71E-02
8	0.0237	0.1047	2.86E-04	0.0520	0.1321	2.12E-03	0.1746	0.2498	4.09E-02
9	0.2392	0.3473	1.54E-03	0.3431	0.4263	1.87E-02	0.4820	0.5116	3.13E-01
10	0.2741	0.3395	2.41E-02	0.3627	0.4190	4.90E-02	0.5087	0.5213	6.16E-01
11	0.3185	0.3149	9.15E-01	0.4567	0.4056	1.25E-01	0.4975	0.5294	2.26E-01
12	0.0133	0.0900	4.02E-05	0.0580	0.1366	1.70E-03	0.1689	0.2995	2.04E-04
13	0.2043	0.1184	6.77E-03	0.3055	0.1369	2.05E-06	0.4810	0.2283	3.19E-10
14	0.2537	0.1812	1.67E-02	0.3270	0.2357	2.92E-03	0.4320	0.3747	5.79E-02
15	0.2377	0.2390	9.63E-01	0.3323	0.3010	3.00E-01	0.4923	0.4195	8.46E-03
16	0.2231	0.2349	6.80E-01	0.2906	0.2920	9.62E-01	0.4431	0.4028	1.43E-01
17	0.1438	0.1956	1.17E-01	0.2353	0.2533	6.32E-01	0.4152	0.3809	3.99E-01
18	0.1629	0.1489	6.29E-01	0.2441	0.1898	9.16E-02	0.4316	0.3235	3.39E-03
19	0.1652	0.2749	2.02E-03	0.2322	0.3291	8.20E-03	0.4322	0.4095	5.38E-01

Based on the significant MI between two EEG channels, a relevance network is derived. For the relevance network, the MI connection with a two-tailed p-value from the ANOVA ($p < 0.001$) is considered significant for the within group analysis (SQ vs. BO) and the across group analysis (N vs. Ppre vs. Ppost). Because we are more interested in the connection with greater MI values, the permutation test is used in conjunction with the ANOVA test to select the relevant features for the MI based relevance network. We have chosen the largest observed value of the permutation test as our threshold to select the significant connections for interpretation.

We first examine the results for the within group analysis (SQ vs. BO). The graphical results at 3 different frequency bands are presented in Fig. 2.11. Since we are comparing the network between different tasks, the subjects are fixed. The solid line denotes that the MIs of the BO task are significantly greater than the ones for the SQ task. The dotted line means that the MIs of the BO task are significantly smaller than the ones for the SQ task. The results suggest that PD subjects are unable to independently recruit different areas of the brain while performing simultaneous tasks, but instead they attempt to recruit a single coherent widespread network. Additionally, medication appears to partially normalize this

deficiency, consistent with prior studies investigating corticomuscular coherence in PD [40].

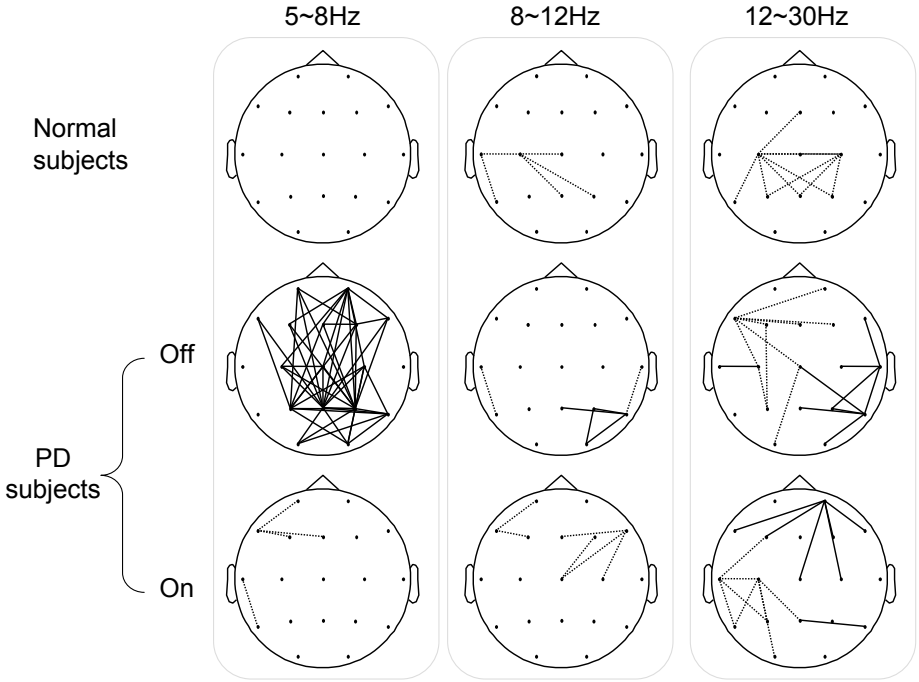


Figure 2.11: Mutual Information based Relevance Network for Within Group Analysis: SQ vs. BO. The solid line denotes that the MIs of the BO task are significantly greater than the ones for the SQ task. The dotted line means the MIs of the SQ task are significantly greater than the ones for the BO task.

We also investigate the results for across subject groups analysis (N vs. Ppre vs. Ppost). In this chapter, we focus mainly on the comparison between N vs. Ppre. Because the comparison is done on different subjects groups, the tasks are fixed. The solid line again represents that the MIs of N are significantly larger than the ones of Ppre. On the other hand, the dotted line shows that the MIs of N are significantly smaller than the ones of Ppre. The results for across group analysis for SQ and BO at 3 different frequency bands are shown in Fig. 2.12. We observe higher MI values in the frontal region at lower and medium frequency bands and motor cortex at higher frequency band in PD which coincides with the finding in the previous graphical theoretical analysis.

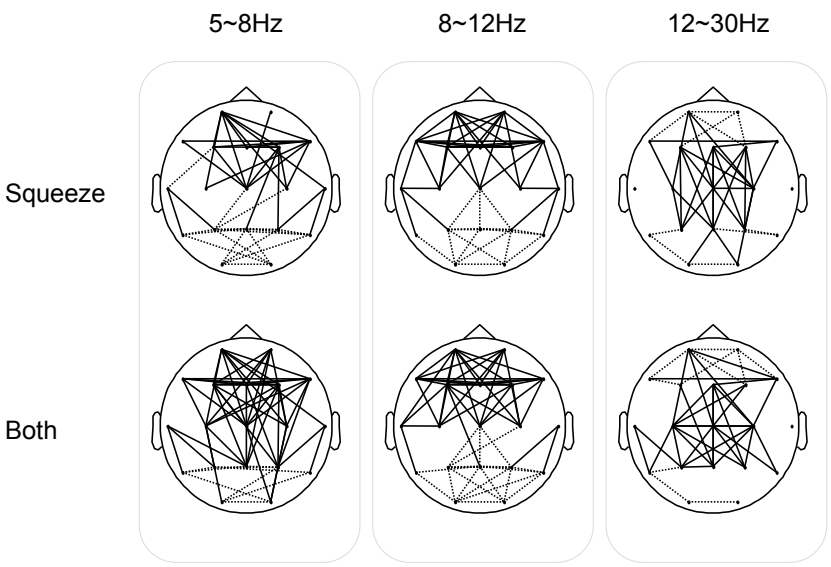


Figure 2.12: Mutual Information based Relevance Network for Across Subject Groups Analysis: N vs. Ppre. The solid line denotes that the MIs of Ppre are significantly greater than the ones for N. The dotted line means the MIs of N are significantly greater than the ones for Ppre.

Furthermore, in order to demonstrate the effectiveness of the EEG segmentation, the relevance networks are also computed for the residual EEG segments that are removed from the analysis. From Fig. 2.13, particularly the Ppost network, the artifacts/noise that spread across the entire brain are being removed by EEG segmentation.

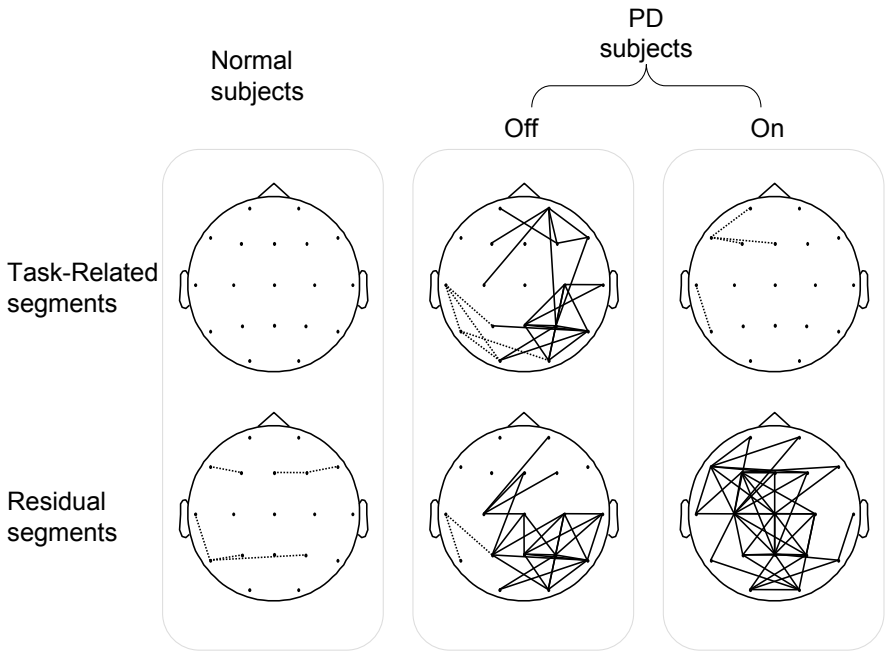


Figure 2.13: An illustration of Mutual Information based Relevance Network for Within Group (SQ vs. BO) based on the task-related EEG segments and the residual EEG segments between 5∼8Hz. The solid line denotes that the MIs of the BO task are significantly greater than the ones for the SQ task. The dotted line means the MIs of the SQ task are significantly greater than the ones for the BO task.

2.4 Conclusion

This chapter proposed a novel segmentation, mutual information network framework for EEG connectivity analysis. Mutual information is able to measure both linear and non-linear dependencies, so it is a suitable metric for identifying brain connectivity networks. EEG segmentation was applied to the data in order to overcome the non-stationarity of the signal and extract the task-related EEG segments for further analysis. A real case study of Parkinson's disease demonstrated that PD subjects have a widespread, excessive synchronization, particularly in the theta (5∼8Hz) band. This excessive synchronization was partially normalized with L-Dopa medication, suggesting a mechanism of action for this medication. Moreover, the results suggest that normal subjects are able to perform simul-

taneous movements by activating different areas of the cortex independently. In contrast, PD subjects must recruit a widespread network in order to perform simultaneous tasks.

Chapter 3

Supplemental Results for Mutual Information Network Analysis

3.1 Introduction

The purpose of this chapter is to expand on the results presented in chapter 2. As chapter 2 mainly focuses on the comparison between normal (N) subjects and Parkinson's disease (PD) subjects - off medication (Ppre), we have added PD - on medication (Ppost) into the comparison here. The methods are described in section 3.2. The results of the graphical theoretical analysis and one-way analysis of variance (ANOVA) test are presented in section 3.3. Section 3.4 concludes the chapter.

3.2 Methods

The motor task involving either PD's right hand only (squeeze task: SQ) or both hands simultaneously (squeeze with right hand and button push with left hand: BO) has been repeated 1 hour after immediate-release formulation of L-Dopa. The collected EEG data are segmented based on the temporal dynamics of the cross-spectrogram of computed Independent Components (ICs). Furthermore, the data are broken down to 4 second epochs for mutual information (MI) computation. The details of the methods are discussed in chapter 2.

3.3 Results

In this section, the results of the graphical theoretical analysis and ANOVA test are investigated. Both the graphical theoretical measure and the ANOVA test are applied to the networks to gain an insight into the effectiveness of the L-Dopa medication on PD subjects.

3.3.1 Graphical Theoretical Analysis

The MI matrices are converted to unidirectional graphs by applying a threshold T. The graphs are characterized by a cluster coefficient C and a mean shortest path length L. C measures the local connectivity of the graph and L measures the global connectivity of the graph. C and L are computed as a function of T ranging from 0.01 to 0.3 with increment of 0.002. We examine the relationship between the graph degree K and T for each group (ie. N-SQ, Ppre-SQ, Ppost-SQ, N-BO, Ppre-BO, and Ppost-BO). Fig. 3.1 shows K as a function of T. They are approximately the same for all cases, because the variances of the mean level of all six groups are very small. Therefore, the rest of the analysis is simply done as a function of T. As T increases, the graphs start to break into subgraphs. At even higher T, the graphs turn into empty graphs. Table 3.1 and 3.2 summarize the points where the graphs start to break into subgraphs or become empty graphs for SQ and BO, respectively. The points are very close in all cases. This also confirms that the mean level of each group does not vary much.

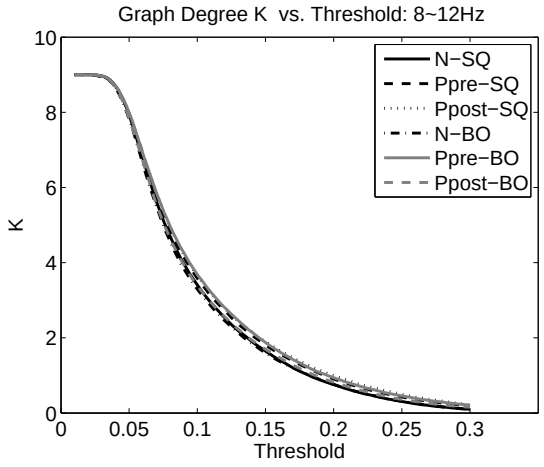


Figure 3.1: Graph degree K as a function of threshold T

Table 3.1: SQ Task: Threshold T when graphs start to split into subgraphs or become empty graphs

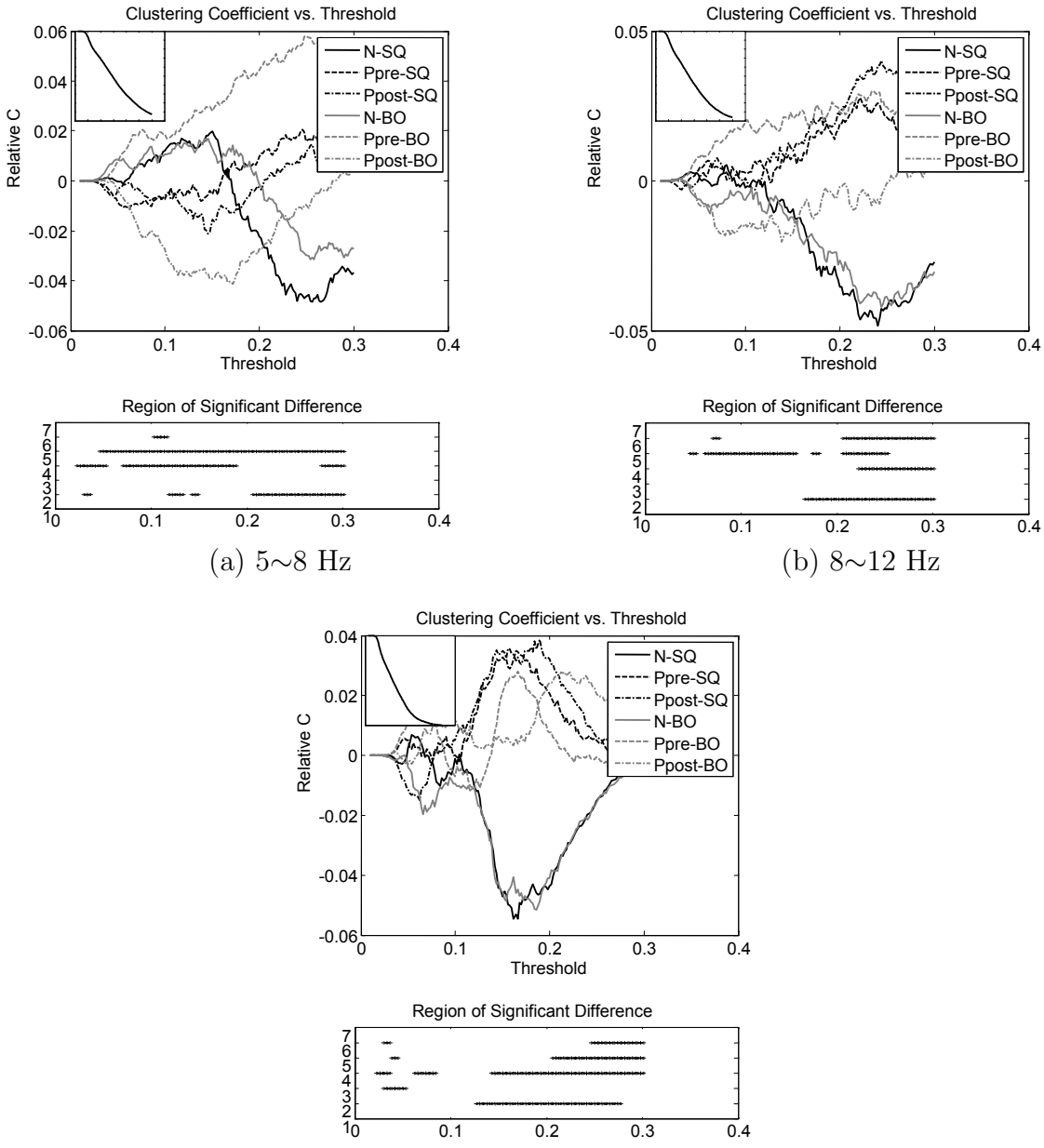
	N-SQ		Ppre-SQ		Ppost-SQ	
	Subgraphs	Empty Graphs	Subgraphs	Empty Graphs	Subgraphs	Empty Graphs
5∼8Hz	0.0600	0.2340	0.0660	0.2380	0.0700	0.2160
8∼12Hz	0.0520	0.2140	0.0560	0.2120	0.0520	0.2060
12∼30Hz	0.0500	0.2060	0.0520	0.1660	0.0520	0.1740

Table 3.2: BO Task: Threshold T when graphs start to split into subgraphs or become empty graphs

	N-BO		Ppre-BO		Ppost-BO	
	Subgraphs	Empty Graphs	Subgraphs	Empty Graphs	Subgraphs	Empty Graphs
5∼8Hz	0.0760	0.2600	0.0600	0.2360	0.0540	0.2220
8∼12Hz	0.0460	0.1980	0.0580	0.1980	0.0640	0.1980
12∼30Hz	0.0440	0.1920	0.0540	0.1640	0.0540	0.1780

The deviation of C from the overall mean of the six groups as a function of T is presented at the top panel of Fig. 3.2. The overall mean of the six groups as a function of T is displayed at the top left corner of the top panel. The bottom panel shows the region that is significantly different between groups (at 2: N vs. Ppost (SQ); at 3: Ppre vs. Ppost (SQ); at 4: N vs. Ppost (BO); at 5: Ppre vs. Ppost (BO); at 6: SQ vs. BO (Ppost)). We observe a general pattern: the cluster intensity for Ppost at BO is reduced compared to the one for Ppre at BO, especially at 5∼8Hz and 8∼12Hz. The consistently high C seen in Ppre reflects the widespread, excessive synchronization seen in PD [39, 41]. The L-Dopa

medication appears to partially normalize the synchronization, consistent with prior studies in the study of corticomuscular coherence in PD [40].



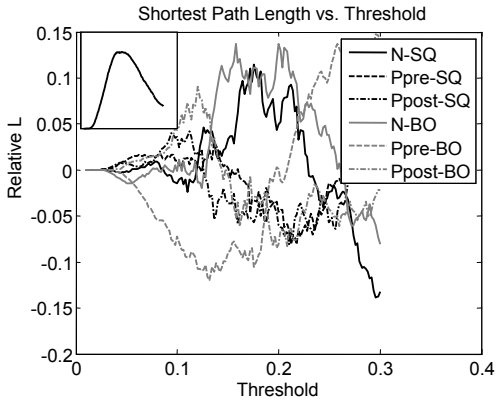
(a) 5∼8 Hz

(b) 8∼12 Hz

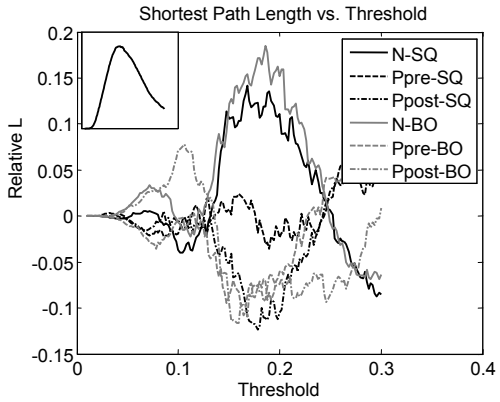
(c) 12∼30 Hz

Figure 3.2: The deviation of C from the overall mean of the six groups as a function of T is shown at the top panel. The overall mean of the six groups as a function of T is displayed at the top left corner of the top panel. The bottom panel shows the region that is significantly different between groups (at 2: N vs. Ppost (SQ); at 3: Ppre vs. Ppost (SQ); at 4: N vs. Ppost (BO); at 5: Ppre vs. Ppost (BO); at 6: SQ vs. BO (Ppost)).

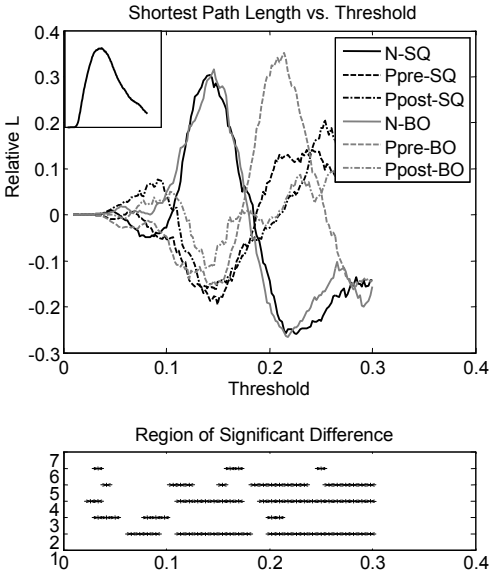
The top panel of Fig. 3.3 shows the deviation of L from the overall mean of the six groups as a function T. The overall mean of the six groups is presented at the top left corner of the top panel. The region that is significantly different between groups (at 2: N vs. Ppost (SQ); at 3: Ppre vs. Ppost (SQ); at 4: N vs. Ppost (BO); at 5: Ppre vs. Ppost (BO); at 6: SQ vs. BO (Ppost)) is shown at the bottom panel. There is no significant difference between L of Ppost at SQ and Ppost at BO. Similarly, Ppre at SQ is quite similar to Ppost at SQ. Here, the effect of medication is not apparent. This might be due to the fact that the graphs start to break into subgraphs which makes it hard to interpret. Nevertheless, N subjects still seem to have the longest L in comparison to the rest of the groups.



(a) 5∼8 Hz



(b) 8∼12 Hz



(c) 12∼30 Hz

Figure 3.3: The deviation of L from the overall mean of the six groups as a function T. The overall mean of the six groups is presented at the top left corner of the top panel. The region that is significantly different between groups (at 2: N vs. Ppost (SQ); at 3: Ppre vs. Ppost (SQ); at 4: N vs. Ppost (BO); at 5: Ppre vs. Ppost (BO); at 6: SQ vs. BO (Ppost)) is shown at the bottom panel.

To study the location of the brain that is more intensely clustered between different groups, the C for each vertex (channel) is summarized in Table 3.3 and 3.4. The significance of C per vertex between a pair of groups is checked by pair-wise T-test. For the SQ task, there is no strong consistency between the locations of the clusters in the brain over all three frequency bands for the groups. However, for the BO case, N subjects have clusters more intensely gathered at the motor cortex region of the brain compared to Ppost subjects, especially at 5∼8Hz and 8∼12Hz. The clusters are also more intensely connected at the frontal region of the brain for Ppre compared to Ppost over all three frequency bands. This observation implies that N subjects' ability to recruit motor cortex region of the brain plays an important role in performing simultaneous tasks properly. However, due to the deficiency in PD, Ppre tries to recruit other regions of the brain (ie. frontal region) in order to help them in achieving the task. The medication has partially normalized the PD subjects.

	5~8Hz					8~12Hz					12~30Hz				
Channel #	N mean	Ppre mean	Ppost mean	P_NPpost	P_PprePpost	N mean	Ppre mean	Ppost mean	P_NPpost	P_PprePpost	N mean	Ppre mean	Ppost mean	P_NPpost	P_PprePpost
1	0.5471	0.5556	0.6279	4.99E-02	4.56E-02	0.6788	0.6752	0.7451	2.82E-02	8.27E-03	0.7395	0.7204	0.7702	1.70E-01	1.46E-02
2	0.4754	0.5054	0.5958	4.01E-03	1.73E-02	0.6229	0.7133	0.6850	6.83E-02	3.39E-01	0.6982	0.7660	0.7146	5.39E-01	2.67E-02
3	0.1520	0.2151	0.2136	7.52E-02	9.68E-01	0.2705	0.3638	0.3305	1.51E-01	4.09E-01	0.3961	0.5643	0.4718	8.40E-02	2.21E-02
4	0.4318	0.5394	0.5180	3.97E-02	5.80E-01	0.6170	0.6470	0.6249	8.14E-01	4.61E-01	0.6781	0.7039	0.6978	4.00E-01	7.86E-01
5	0.4001	0.5409	0.4839	8.61E-03	4.80E-02	0.5648	0.6767	0.5878	3.84E-01	2.88E-04	0.6439	0.7440	0.6607	4.27E-01	4.16E-05
6	0.4471	0.6156	0.5790	1.39E-03	3.11E-01	0.5363	0.6728	0.6078	2.31E-02	1.68E-02	0.6148	0.7013	0.6313	4.96E-01	1.55E-03
7	0.1368	0.1993	0.1512	6.69E-01	1.28E-01	0.2806	0.3679	0.3710	3.89E-02	9.39E-01	0.4731	0.4819	0.5363	1.52E-01	1.59E-01
8	0.0076	0.0492	0.0036	5.46E-01	1.33E-03	0.0312	0.1073	0.0551	1.97E-01	2.51E-02	0.1767	0.2333	0.1530	4.85E-01	1.44E-02
9	0.1286	0.2163	0.1775	1.21E-01	2.22E-01	0.3395	0.3392	0.2743	6.26E-02	4.87E-02	0.5177	0.4640	0.4113	6.81E-04	8.41E-02
10	0.1144	0.2434	0.1975	2.23E-03	1.31E-01	0.3466	0.3913	0.3403	8.39E-01	9.36E-02	0.5209	0.5234	0.4954	3.25E-01	2.85E-01
11	0.1462	0.2433	0.1976	1.20E-01	1.67E-01	0.3593	0.4065	0.3328	4.45E-01	2.42E-02	0.5156	0.5022	0.4381	8.93E-03	1.92E-02
12	0.0000	0.0337	0.0390	5.30E-03	7.45E-01	0.0453	0.1346	0.1214	3.52E-03	6.42E-01	0.1734	0.2746	0.2657	1.29E-02	7.98E-01
13	0.0830	0.0913	0.0128	4.05E-04	7.20E-05	0.2824	0.1384	0.0869	4.66E-08	6.24E-02	0.4799	0.2557	0.2099	2.98E-10	2.00E-01
14	0.1554	0.1245	0.1366	5.34E-01	6.51E-01	0.3117	0.2322	0.2101	1.26E-03	4.60E-01	0.4765	0.3414	0.3413	1.21E-05	9.99E-01
15	0.1614	0.1312	0.1615	9.99E-01	2.75E-01	0.3340	0.2310	0.2752	7.65E-02	1.19E-01	0.5111	0.3665	0.4020	5.40E-04	2.35E-01
16	0.1395	0.1289	0.1934	8.53E-02	2.37E-02	0.2691	0.2423	0.2685	9.83E-01	3.54E-01	0.4560	0.3563	0.3909	2.13E-02	2.06E-01
17	0.0589	0.0746	0.1156	3.77E-02	1.09E-01	0.2228	0.1265	0.2727	2.19E-01	1.07E-05	0.4521	0.2778	0.4685	7.16E-01	1.07E-06
18	0.0679	0.1088	0.0278	2.45E-02	3.43E-04	0.2029	0.1955	0.1330	2.03E-02	4.99E-02	0.4547	0.3436	0.2965	4.41E-05	2.03E-01
19	0.0752	0.0843	0.1111	1.82E-01	3.00E-01	0.2083	0.1900	0.2510	2.45E-01	7.57E-02	0.4615	0.3422	0.4118	2.12E-01	5.95E-02

Table 3.3: SQ Task: C for each vertex and the p-value of the pair-wise T-test between groups. (T = 0.22 for 5∼8Hz; T = 0.18 for 8∼12Hz; T =0.15 for 12∼30Hz) An example of notation convention: P_NPpost = P-value for N vs. Ppost.

	5∼8Hz					8∼12Hz					12∼30Hz				
Channel #	N mean	Ppre mean	Ppost mean	P_NPpost	P_PprePpost	N mean	Ppre mean	Ppost mean	P_NPpost	P_PprePpost	N mean	Ppre mean	Ppost mean	P_NPpost	P_PprePpost
1	0.5751	0.5719	0.5040	9.82E-02	7.55E-02	0.7140	0.7028	0.6752	2.36E-01	3.17E-01	0.7743	0.7636	0.7213	2.58E-02	3.63E-02
2	0.4989	0.5241	0.5022	9.38E-01	5.75E-01	0.6801	0.7089	0.6638	6.28E-01	1.31E-01	0.7158	0.7630	0.7112	8.56E-01	1.67E-02
3	0.1413	0.2417	0.1028	1.67E-01	1.75E-05	0.2286	0.3980	0.2272	9.70E-01	1.10E-05	0.3552	0.5811	0.4413	4.51E-02	4.73E-04
4	0.4927	0.6032	0.4733	6.46E-01	6.27E-04	0.6233	0.6993	0.5985	4.64E-01	1.05E-03	0.6844	0.7434	0.6432	1.05E-01	1.52E-05
5	0.4536	0.5296	0.4481	8.71E-01	9.14E-03	0.5661	0.6955	0.5365	2.52E-01	1.50E-10	0.6334	0.7611	0.6267	7.52E-01	4.14E-11
6	0.4043	0.5496	0.4841	4.03E-02	6.86E-02	0.5972	0.6808	0.6015	8.89E-01	4.65E-03	0.6668	0.7204	0.6082	1.27E-02	7.13E-08
7	0.1415	0.2391	0.1098	2.93E-01	3.46E-05	0.3399	0.4300	0.2746	1.11E-01	3.16E-05	0.5210	0.6103	0.4608	1.47E-01	3.53E-05
8	0.0083	0.0637	0.0045	6.14E-01	1.56E-04	0.0520	0.1321	0.0346	3.44E-01	4.84E-05	0.1746	0.2498	0.0874	3.72E-03	4.55E-07
9	0.1500	0.2743	0.1778	3.83E-01	2.99E-03	0.3431	0.4263	0.2961	1.86E-01	1.68E-04	0.4820	0.5116	0.3984	7.76E-03	2.27E-04
10	0.1730	0.2792	0.1914	5.31E-01	2.21E-03	0.3627	0.4190	0.3063	5.09E-02	1.25E-04	0.5087	0.5213	0.4337	5.09E-03	1.08E-03
11	0.1746	0.2567	0.1726	9.48E-01	3.89E-03	0.4567	0.4056	0.2804	2.76E-07	4.51E-05	0.4975	0.5294	0.4322	2.21E-02	3.91E-04
12	0.0000	0.0787	0.0721	1.86E-05	7.63E-01	0.0580	0.1366	0.1392	1.68E-03	9.20E-01	0.1689	0.2995	0.2542	1.68E-02	1.93E-01
13	0.1302	0.1057	0.0333	1.97E-04	7.22E-04	0.3055	0.1369	0.0983	2.16E-08	1.55E-01	0.4810	0.2283	0.2436	5.07E-08	6.67E-01
14	0.1771	0.1416	0.0868	9.48E-04	2.21E-02	0.3270	0.2357	0.1897	1.10E-05	1.20E-01	0.4320	0.3747	0.3588	1.83E-02	6.11E-01
15	0.1876	0.1839	0.1562	3.08E-01	3.13E-01	0.3323	0.3010	0.2488	6.89E-03	5.85E-02	0.4923	0.4195	0.3901	5.60E-04	2.79E-01
16	0.1618	0.1827	0.1667	8.62E-01	5.27E-01	0.2906	0.2920	0.2835	8.18E-01	7.61E-01	0.4431	0.4028	0.3859	5.06E-02	5.37E-01
17	0.0982	0.1790	0.1820	1.04E-02	9.29E-01	0.2353	0.2533	0.2700	3.88E-01	6.50E-01	0.4152	0.3809	0.4163	9.80E-01	3.68E-01
18	0.1227	0.1339	0.0594	9.43E-03	3.37E-03	0.2441	0.1898	0.1585	7.83E-03	3.12E-01	0.4316	0.3235	0.3072	1.02E-03	6.45E-01
19	0.1217	0.2269	0.2179	5.44E-03	8.02E-01	0.2322	0.3291	0.2891	1.29E-01	2.76E-01	0.4322	0.4095	0.3949	3.45E-01	6.88E-01

Table 3.4: BO Task: C for each vertex and the p-value of the pair-wise T-test between groups. (T = 0.22 for 5∼8Hz; T = 0.18 for 8∼12Hz; T =0.15 for 12∼30Hz) An example of notation convention: P_NPpost = P-value for N vs. Ppost.

3.3.2 One-way ANOVA Test

Across group differences of each MI value are analyzed using ANOVA with a subject group factor. The connection is established if it is significantly different (ie. $p < 0.001$) between the two subject groups and the mean of the connection is above the largest observed value of the permutation test. The significant connections are shown in the relevance network. Fig. 3.4 gives the networks for N vs. Ppost and Ppre vs. Ppost while they are doing the SQ task. The networks for the same subject groups doing the BO task are shown in Fig. 3.5.

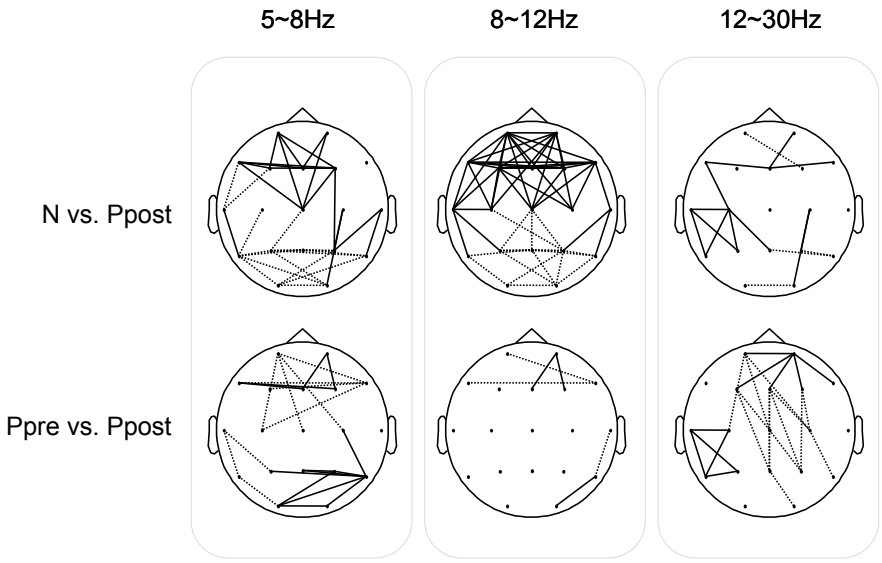


Figure 3.4: Mutual Information based Relevance Network for Across Subject Groups Analysis: SQ task. For the top row (N vs. Ppost), the solid line denotes that the MIs for Ppost are significantly greater than the ones for N. The dotted line means the MIs for N are significantly greater than the ones for Ppost. For the bottom row (Ppre vs. Ppost), the solid line denotes that the MIs for Ppost are significantly greater than the ones for Ppre. The dotted line means the MIs for Ppre are significantly greater than the ones for Ppost.

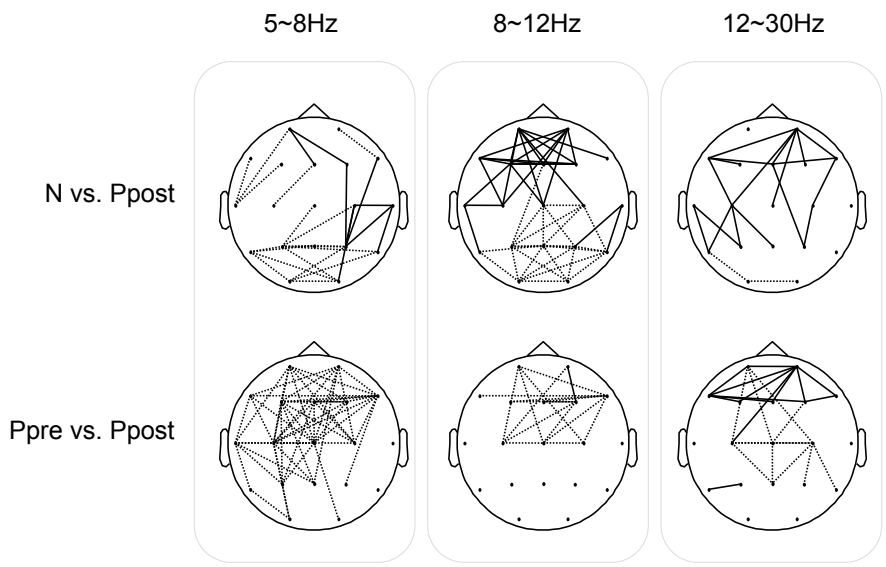


Figure 3.5: Mutual Information based Relevance Network for Across Subject Groups Analysis: BO task. For the top row (N vs. Ppost), the solid line denotes that the MIs for Ppost are significantly greater than the ones for N. The dotted line means the MIs for N are significantly greater than the ones for Ppost. For the bottom row (Ppre vs. Ppost), the solid line denotes that the MIs for Ppost are significantly greater than the ones for Ppre. The dotted line means the MIs for Ppre are significantly greater than the ones for Ppost.

For the SQ task, the networks for Ppre and Ppost do not differ as much between 5∼8Hz and 8∼12Hz. However, at 12∼30Hz, there is a significant reduction of MI values for Ppost at the motor cortex region of the brain which results in a smaller network difference between N and Ppost. This suggests that the medication normalizes the brain functioning of PD subjects while they are doing the SQ task at high frequency band. For the BO task, after the PD subjects have taken the medication, there is a significant reduction of the MI values at the frontal region and motor cortex of the brain at 5∼8Hz and at the frontal region at 8∼12Hz. This also suggests to us that L-Dopa medication is normalizing the PD toward N. From the ANOVA test, we can conclude that L-Dopa medication makes more a difference for PD subjects when they are performing the BO task. This finding agrees with the physiological observation. PD subjects particularly have difficulties executing simultaneous movements. However, after the release of L-Dopa, their performance of simultaneous movements are improved [40].

3.4 Conclusion

We proposed a novel EEG segmentation method and applied MI network analysis in the study of PD. EEG segmentation allows us to obtain the quasi-stationary epochs from the data for our MI analysis. MI measures the dependencies between two EEG channels, so that we can derive a functional connectivity network. Both the graphical theoretical measure and the ANOVA test demonstrated statistically significant differences between each group. A real case study of PD demonstrated the effect of L-Dopa medication. The medication partially normalizes the PD subjects and improves the PD's ability in performing simultaneous movements.

Chapter 4

Conclusion

Chapter 4 presents the conclusion and contribution and outlines recommendations for future research arising from the study. The chapter includes an overview, restatement of the research problems, review of methodology, review of the findings and the possible future work.

4.1 Conclusion and Contribution

In this research, we presented a mutual information network approach to analyze multi-channel EEG data collected from normal and Parkinson's disease subjects while they were performing motor tasks. Accessing the functional connectivity networks of the brain has gained increasing interest because the networks help the researchers in understanding the brain, particularly in disease states. Current methods of detecting networks in the recorded EEG are limited by the fact that they assume stationarity of the relationship between channels, and their reliance on linear dependencies. In order to overcome these problems,

- We derived a novel segmentation method based on the temporal dynamics of the cross-spectrogram of the computed independent components. The proposed non-parametric segmentation approach does not require a priori information about the underlying parameters in the signals. After suitable thresholding of the cross-spectrogram values of the computed independent components, the quasi-stationary and task-related segments of the EEG were obtained. (Chapter 2)

- We applied mutual information to determine the nonlinear statistical dependencies

between EEG channels and derived the relevance network. Mutual information between two EEG channels was computed based on the segmented data. Since mutual information quantifies the information transmitted from one signal to another, it can thus detect the functional coupling in the brain. By estimating the mutual information between the time series of multiple pairs of EEG channels, a network can be derived. (Chapter 2)

The analysis was applied to the study of Parkinson's disease data. Because Parkinson's disease is a motor disorder, the subjects exhibit difficulties in carrying out simultaneous movements. EEG recordings are obtained from normal subjects, Parkinson's subjects - off medication, and Parkinson's subjects - on medication while they are performing a motor task either by one hand or both hands. In chapter 2 and chapter 3, graphical theoretical measure and one-way analysis of variance test were applied to analyze the network differences. The networks are statistically significantly different between groups. We found

- Parkinson's subjects have a widespread, excessive synchronization in the brain.
- The excessive synchronization is partially normalized with L-Dopa medication, suggesting a mechanism of action for this mediation.

4.2 Future Work

The mutual information analysis here investigated the pair-wise, unidirectional relationship between EEG channels. The current method can be further improved in a number of directions:

- To explore the causality relationship between EEG channels by investigating the transfer entropy. Transfer entropy measures how the transition probabilities of one signal influence the other and can give information about the direction of interaction.

- To expand from bivariate to multivariate so that the relationship between three or more EEG channels can be studied.
- To examine the interdependencies between the independent components of the data, which has been suggested by a few recent studies.

Bibliography

[1] R. Spehlmann, *EEG Primer*. New York, NY: Elsevier North-Holland, 1981.

[2] F. S. Tyner, J. R. Knott, and W. B. Mayer, *Fundamentals of EEG Technology: Vol. 2 Clinical Correlates*. New York, NY: Raven Press, 1983.

[3] C. J. Stam, *Nonlinear Brain Dynamics*. New York, NY: Nova Science Publishers, 2006.

[4] J. Jeong, J. C. Gore, and B. S. Peterson, "Mutual information analysis of the eeg in patients with alzheimer's disease," *Clinical Neurophysiology*, vol. 112, pp. 827–835, 2001.

[5] S. H. Na *et al.*, "Eeg in schizophrenic patients: mutual information analysis," *Clinical Neurophysiology*, vol. 113, pp. 1954–1960, 2002.

[6] A. Y. Kaplan *et al.*, "Nonstationary nature of the brain activity as revealed by eeg/meg: Metholdological, practical and conceptual challenges," *Signal Processing*, vol. 85, pp. 2190–2212, 2005.

[7] F. H. L. da Silva, *Analysis of EEG Non-Stationarities. In W.A. Cobb & H. Van Duijn, Contemporary clinical neurophysiology*. Amsterdam: Elsevier.

[8] B. E. Brodsky and B. S. Darkhovsky, *Nonparametric Statistcal Diagnosis: Problems and Methods*. Dordrecht (the Netherlands):Kluwer Academic Publishers, 2000.

[9] K. A. KOOI, R. P. Tucker, and R. E. Marshall, *Fundamentals of electroencephalography*. Hagerstown, MD: Medical Department, Harper & Row, 1978.

[10] J. S. Ebersole and T. A. Pedley, *Current Practice of Clinical Electroencephalography*. Philadelphia, PA: Lippincott Williams & Wilkins, 2003.

[11] K. E. Misulis and T. C. Head, *Essentials of Clinical Neurophysiology*. Boston, MA: Butterworth-Heinemann, 2003.

[12] E. Perda, R. Quiroga, and J. Bhattacharya, "Nonlinear multivariate analysis of neurophysiological signsl," *Progress in Neurobiology*, vol. 77, pp. 1–37, 2005.

[13] R. Q. Quiroga, J. Arnhold, and P. Grassberger, "Learning driver-response relationships from synchronization patterns," *Physical Review E*, vol. 61, no. 5, pp. 5149–5153, 2000.

[14] F. L. da Silva, J. Pijn, and P. Boeijinga, "Interdependence of eeg signals: linear vs. nonlinear associations and the significance of time delays and phase shifts," *Brain Topography*, vol. 2, pp. 9–18, 1989.

[15] C. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, pp. 424–438, 1969.

[16] T. Tersvirta, "Modelling economic relationships with smooth transition regressions," Stockholm School of Economics, Working Paper Series in Economics and Finance 131, 1996.

[17] W. Freiwald *et al.*, "Testing non-linearity and directedness of interactions between neural groups in the macaque inferotemporal cortex," *Journal of Neuroscience Methods*, vol. 94, pp. 105–119, 1999.

[18] B. Shen, M. Nadkarni, and R. A. Zappulla, "Spectral modulation of cortical connections measured by eeg coherence in human," *Clinical Neurophysiology*, vol. 110, pp. 115–125, 1999.

[19] S. Weiss and P. Rappelsberger, "Long-range eeg synchronisation during word encoding correlates with successful memory performance," *Cognitive Brain Ressearch*, vol. 9, pp. 299–312, 2000.

[20] B. C. Min *et al.*, "Analysis of mutual information content for eeg responses to odor stimulation for subjects classified by occupation," *Chem. Senses*, vol. 28, pp. 741–749, 2003.

[21] R. Benecke *et al.*, "Performance of simultaneous movements in patients with parkinson's disease," *Brain*, vol. 109, pp. 793–757, 1986.

[22] A. Berardelli *et al.*, "Pathophysiology of bradykinesia in parkinson's disease," *Brain*, vol. 124, pp. 2131–2146, 2001.

[23] O. Sporns and J. D. Zwi, "The small world of the cerebral cortex," *Neuroinformatics*, vol. 2, pp. 145–162, 2004.

[24] S. H. Strogatz, "Exploring complex networks," *Nature*, vol. 410, pp. 268–276, 2001.

[25] C. J. Stam *et al.*, "Small-world networks and functional connectivity in alzhiemer's disease," *Cerebral Cortex*, vol. 17, no. 1, pp. 92–99, 2007.

[26] G. Pfurtscheller and F. H. L. da Silva, "Event-related eeg/meg synchronization and desynchronization: basic principles," *Clinical Neurophysiology*, vol. 110, pp. 1842–1857, 1999.

[27] T.-P. Jung *et al.*, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, vol. 37, pp. 163–178, 2000.

[28] A. Bell and T. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Computation*, vol. 7, pp. 1129–1159, 1995.

[29] A. Delorme and S. Makeig, "Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 134, pp. 9–21, 2004.

[30] A. Delorme *et al.*, "From single-trials eeg to brain area dynamics," *Neurocomputing*, vol. 46, p. 10571064, 2002.

[31] D. Gupta and C. J. James, "Narrowband vs. broadband phase synchronization analysis applied to independent components of ictal and interictal eeg," in *Proc. 29th Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBS29)*, 2007.

[32] B. Hong *et al.*, "Transient phase synchrony of independent cognitive components underlying scalp eeg," in *Proc. 27th Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBS27)*, 2005.

[33] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Upper Saddle River, NJ: Prentice-Hall, 1999.

[34] T. M. Cover and J. A. Thomas, *Elements of information theory*. New York: John Willey & Sons, 1991.

[35] A. J. BUTTE and I. S. KOHANE, "Mutual information relevance networks: functional genomic clustering using pairwise entropy measurements," *Pacific Symposium on Biocomputing*, vol. 5, pp. 415–426, 2000.

[36] R. V. Hogg and J. Ledolter, *Engineering statistics*. New York: MacMillan, 1987.

[37] D. Francois, V. Wertz, and M. Verleysen, "The permutation test for feature selection by mutual information," in *Proc. 14th European Symposium on Artificial Neural Networks (ESANN14)*, 2006.

[38] M. Hoehn and M. Yahr, "Parkinsonism: onset, progression and mortality," *Neurology*, vol. 17, pp. 427–442, 1967.

[39] L. M. Doyle *et al.*, "Levodopa-induced modulation of subthalamic beta oscillations during self-paced movements in patients with parkinson's disease," *European Journal of Neuroscience*, vol. 21, pp. 1403–1412, 2005.

[40] S. Salenius *et al.*, "Defective cortical drive to muscle in parkinsons disease and its improvement with levodopa," *Brain*, vol. 125, pp. 491–500, 2002.

[41] L. Timmermann *et al.*, "The cerebral oscillatory network of parkinsonian resting tremor," *Brain*, vol. 126, pp. 199–212, 2003.

[42] A. Castanon *et al.*, "Region of interest based analysis of functional imaging data," *NeuroImage*, vol. 19, pp. 1303–1316, 2003.

[43] Y. Liu *et al.*, "Temporal dissociation of parallel processing in the human subcortical outputs," *Nature*, vol. 400, pp. 364–367, 1999.

[44] K. M. Petersson *et al.*, "Statistical limitations in functional neuroimaging. i. non-inferential methods and statistical models," *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, vol. 354, pp. 1239–1260, 1999a.

[45] ——, "Statistical limitations in functional neuroimaging. ii. signal detection and statistical inference," *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, vol. 354, pp. 1261–1281, 1999b.

[46] L. Shi, P. A. Heng, and T.-T. Wong, "A spectral clustering approach to fmri activation cetection," in *Proc. 27th IEEE Int. Conf. IEEE Engineering in Medicine and Biology Society (EMBS27)*, 2005.

[47] M. Meila and S. J., "A random walks view of spectral segmentation," in *Proc. 8th Int. Workshop on Artificial Intelligence and Statistics (AISTATS08)*, 2001.

[48] J. Marchini and B. Ripley, "A new statistical approach to detecting significant activation in functional mri," *NeuroImage*, vol. 12, no. 4, 2000.

[49] P. Mitra and B. Pesaran, "Analysis of dynamic brain imaging data," *Biophysical Journal*, vol. 76, no. 2, 1999.

[50] R. Baumgartner, C. Windischberger, and E. Moser, "Quantification in functional magnetic resonance imaging: fuzzy clustering vs. correlation analysis," *Magnetic Resonance Imaging*, vol. 16, no. 2, 1998.

[51] X. Golay *et al.*, "A new correlation-based fuzzy logic clustering algorithm for fmri," *Magnetic Resonance in Medicine*, vol. 40, no. 2, pp. 249–260, 1998.

[52] F. G. Meyer and J. Chinrungrueng, "Spatiotemporal clustering of fmri time series in the spectral domain," *Medical Image Analysis*, vol. 9, no. 1, pp. 51–68, 2005.

[53] G. A. F. Seber and A. J. Lee, *Linear Regression Analysis*. New York: Wiley, 2003.

[54] C. Buhmann *et al.*, "Pharmacologically modulated fmri-cortical responsiveness to levodopa in drug-naive hemiparkinsonian patients," *Brain*, vol. 126, no. 2, pp. 451–461, 2003.

Appendix A

Spectral Clustering of fMRI Data within Regions of Interest [2]

A.1 Introduction

One common approach for determining task-related activity across subjects in fMRI data is to draw Regions of Interest (ROIs) individually on each subject, and then look at the statistical properties of the same ROI across subjects. Because the ROI approach has been suggested to offer finer localization and increased sensitivity to task related effects, we consider manually segmented anatomically-defined ROIs in this study [42].

To address the question of whether an ROI is activated or not, a simple way would be to calculate the average intensity over an ROI at every time point, and determine if the resulting average time course significantly correlates with the stimulus [43], though this risks including voxels that may not be active, or even voxels that may have been inadvertently included in the ROI, reducing the effect size. An alternative approach is to define a "functional ROI" as the intersection between an anatomically-defined ROI and voxels deemed task-related to a certain degree based on a voxel-level analysis.

In order to identify if ROIs are task-related, classification and clustering techniques can be used [44, 45]. However, the disadvantage of a classification method is that it requires a training data set which is difficult to obtain for real fMRI data. As well, the existing

[2] A version of this chapter has been published.

clustering methods assume that the structure of the data conforms to certain shapes in the features space. This is problematic in fMRI data which has complex statistical properties, and is not typically Gaussian. Spectral clustering overcomes these problems, since it does not require a training data set, and is able to detect clusters of arbitrary shapes according to the natural data structure [46]. Spectral clustering relies on a similarity function between pairs of voxels and optimizes the criterion that voxels belong to the same cluster are similar, whereas voxels in different clusters are dissimilar. This method uses eigenvalues and eigenvector of the similarity matrix, and its computational efficiency is achieved using sparse matrix techniques [47].

In this chapter, we propose to apply spectral clustering on the fMRI data in a Parkinson's disease (PD) study for activation detection and for identifying different activation patterns between normal subjects, and PD subjects before medication and after L-Dopa medication. The proposed approach is different in two ways from the one in [46], where spectral clustering is performed on the Discrete Wavelet Transform (DWT) coefficients and the whole brain. First, different features and similarity functions are explored by taking into consideration our specific fMRI study. Second, group comparison is further studied to examine the effects of disease and medicine.

The chapter is organized as follows: section A.2 describes the activation detection method. The experiments and results on both synthetic fMRI data and the real fMRI data are presented in section A.3 and A.4. We then conclude the chapter.

A.2 Methods

A.2.1 Preprocessing

After standard preprocessing operations (e.g. motion correction, smoothing, and alignment), the fMRI data are detrended to remove any linear trends. Since the stimulus for our experiment is periodic, utilizing a block-design, clustering can be applied to band-limited

data in order to eliminate the irrelevant frequency ranges and reduce the noise. We examine the power spectrum of the fMRI data to select the frequency range of interest [48, 49]. Fig. A.1 gives an example of the power spectral of the mean of an active cluster and a non-active cluster. The variance of the power spectrum over the active voxels is greater than the variance over the non-active voxels [50]. If an ROI contains a mixture of active and non-active voxels, its variance of the power spectrum over the entire region will be greatest at the stimulus frequency. Therefore, only the frequencies with the greatest variance are kept for further analysis. The fMRI data in the experiment are filtered to between 0.0098Hz and 0.0488Hz.

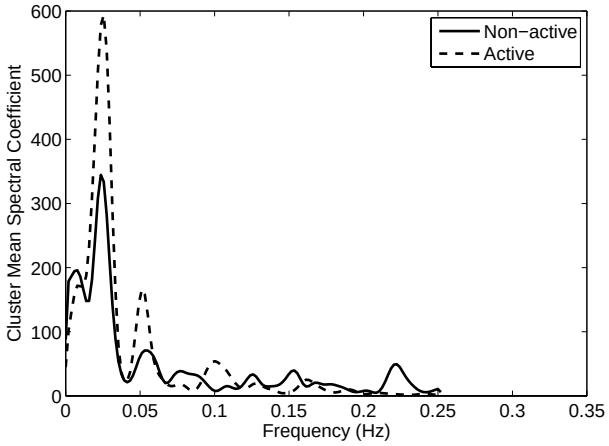


Figure A.1: Spectral coefficients of the mean of two clusters obtained by spectral clustering. The active cluster mean has a sharper peak at the frequency of interest.

A.2.2 Feature Extraction

Euclidean distance, correlation coefficient, and the two combined are indexes frequently used for spectral clustering. Euclidean distance measures the distance between two points. In this case, the Euclidean distance captures the distance between the magnitude of two voxels. The shorter the distance, the greater the similarity and vice versa. Correlation coefficient measures the strength and direction of a linear relationship between the times-series of two voxels. The closer it is to 1, the greater the similarity. In order to combine Euclidean

distance to the correlation coefficient, a transformation, g , must be performed, where g is a continuously and monotonically decreasing function. The goal of g is to translate a large correlation ($\rho \approx 1$) to a smaller distance ($d_c \approx 0$). In order to penalize $\rho \approx -1$, a non-linear function is used [51],

$$g(x) = (\frac{1-x}{1+x})^{\beta}. \tag{A.1}$$

In [46], it also suggests to raise the index to exponential to increase the difference between similarity and dissimilarity. Therefore, we look at these three indexes:

1. Euclidean distance index, $e^{\frac{-||de(i)-de(j)||}{\sigma_e^2}}$
2. Correlation index, $e^{\frac{-||g(i)-g(j)||}{\sigma_c^2}}$
3. The combined index, $e^{\frac{-||de(i)-de(j)||}{\sigma_e^2}} * e^{\frac{-||g(i)-g(j)||}{\sigma_c^2}}$.

A.2.3 Spectral Clustering

Spectral clustering uses eigenvalues and eigenvectors of a matrix constructed from the pairwise similarity function of the data set. For each pair of voxels, $i, j \in I$, a similarity, $S_{ij} = S_{ij} \geq 0$, is defined. The pairwise similarities are viewed as edge flows in a Markov random walk. The properties of the eigenvalues and eigenvectors of the similarity matrix are used for clustering. We apply Meila-Shis Multicut Algorithm in [47], as follows.

1. Compute P from S as

$$P = D^{-1}, \tag{A.2}$$

   where D is a diagonal matrix of $d_i = d_i = \sum_{j \in I} S_{ij}$.

2. Compute $v_1, v_2, \ldots, v_k$, the eigenvectors of P corresponding to the K largest eigenvalues.

3. Form $V_{nxk} = [v_1, v_2, \ldots, v_k]$ with the vectors as column vectors.

4. Cluster the rows of V as voxels in a K-dimensional space.

From the theory of Markov random walks, P_{ij} in A.2 represents the probability of moving from node i to j in one step, given that we are in i . The last step is achieved by k-means in the $(k-1)$ dimensional space defined by the rows of $[v_2, \ldots, v_k]$.

A.3 Experiments with Synthetic Data

A.3.1 Synthetic Data

In order to compare spectral clustering to k-means and also compare different similarity indexes, synthetic data which give us the ground truth are used. Similar to [52], a three-dimensional fMRI data set where a small region is activated can be generated. The MR signal of voxel v at time t is the magnitude of the sum of a complex signal $A_v(t)e^{i\theta(t)}$ and a centered complex Gaussian white noise $n^c(t)$ with unit variance,

$$s_v = |A_v(t)e^{i\theta(t)} + n_v^c(t)|. \tag{A.3}$$

The amplitude A_v is a sinusoidal function of time,

$$A_v = M_v \sin(w_s t + \varphi_v). \tag{A.4}$$

If v is a non-active voxel, A_v is set to 0. The level of activation, M_v , is varied from 0 to 10 in order to see how each method or index performs. Without loss of generality, $\theta_v(t)$ is chose to be $\pi/4$ (the real and imaginary channels play a symmetric role). Discrete time samples, T , is 80, and the frequency of the signal is $w_s = \pi/10$. The random delay φ_v is Gaussian distributed with zero mean and unit variance.

A.3.2 Results

Fig. A.2 shows the true activation detection rate of k-means clustering and spectral clustering using different similarity indexes as the activation level, M_v , varies. The detection rate is the ratio between the number of voxels correctly identified and the total number of voxels. It is clear that the spectral clustering approaches perform better than k-means as illustrated in the figure. Correlation index seems to perform better in this case. After running some tests on our real fMRI data, we noted that correlation index also gives the best overall results. Therefore, only the correlation index is considered for the real fMRI data.

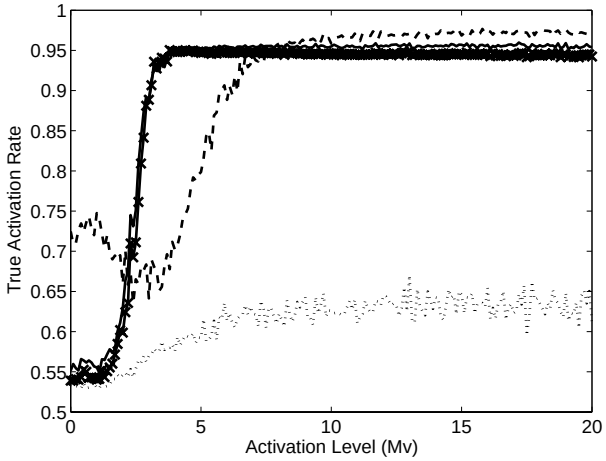


Figure A.2: The true activation rate for different methods and similarity indexes as the activation level varies (SC index 1 - spectral clustering with Euclidean index, SC index 2 - spectral clustering with correlation index, SC index 3 - spectral clustering with combined index of Euclidean and correlation).

A.4 Experiments with fMRI Data

A.4.1 Real fMRI Data

A pressure-responsive squeeze-bulb is used as a response device in our fMRI studies, allowing us to create tasks which require continuous force output at specific amplitudes, times and

speeds. In the experiment, subjects are asked to perform continuous task using a block design with period of constant force at three frequencies: 0.25, 0.5, and 0.75Hz, with target force output scaled to between 5 and 15% MVC. The task is in the order of six sessions: high, low, medium, low, high and medium. A 20-second block design with subjects squeezing in three frequencies is recorded, followed by a resting period of 23 seconds. The fMRI data are obtained from six normal subjects and five Parkinson's disease subjects both before and after medication. Eighteen ROIs were defined and manually traced for each subject based on anatomical sulcal landmarks and with the guidance of a brain atlas. Subjects were tested before and $\sim$ 1hr after given 100mg of L-Dopa.

A.4.2 Results

The illustrations are based on left hand performance. Each ROI is clustered into two clusters corresponding to a more active cluster and a less active (or non-active) cluster. As one example illustrated in Fig. A.3, for an ROI that is known to be active, the proposed spectral clustering is able to separate them into an active cluster and a non-active one. Fig. A.4 gives an example of the active regions in the brain detected by spectral clustering. The observed activation patterns are well-supported by medical knowledge. For further group analysis, robust linear regression [53] is used to fit the time-series of each cluster mean as a function of the three input stimulus box designs and the coefficients are further analyzed for group comparison. The significance of the estimated coefficients within a group indicates the activation level. The active cluster is used to represent the ROI activity. The ROIs that appear as significantly activated in at least one group are summarized in Table A.1, where the corresponding p-values of the group activities are reported. In the table, "N" means the control group; "PA" means the PD group before medication; "PB" means the PD group after medication. The results demonstrate activity in the contralateral motor cortex R_M1, but also suggest a role of the ipsilateral motor cortex (L_M1) in PD at medium and higher task frequencies.

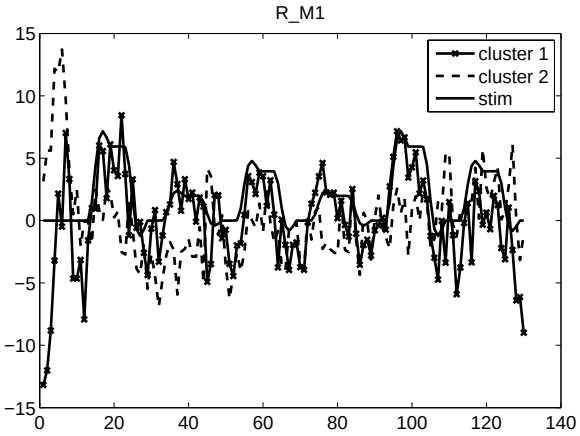


Figure A.3: Cluster representations of the ROI R_M1 (Right Motor), which is known to be task-related (active). Cluster 1 (active) is very similar to the stimulus compares to Cluster 2 (non-active).

Figure A.4: The active voxels detected by spectral clustering. The white regions are the active regions. The black regions are the no-active regions.

Table A.1: P-value of the Activation Level within each Group

ROIs	F1=0.25Hz			F2=0.5Hz			F3 = 0.75Hz		
	N	PA	PB	N	PA	PB	N	PA	PB
L_ACC	0.2968	0.7439	0.6763	0.0085	0.4903	0.0874	0.0006	0.3479	0.5750
L_CAU	0.1867	0.8656	0.8223	0.1788	0.5275	0.0018	0.0672	0.2980	0.3344
L_CER	0.3174	0.0510	0.6527	0.0513	0.3219	0.0165	0.1752	0.0015	0.0358
L_M1	0.4226	0.3618	0.3344	0.4194	0.0026	0.0062	0.0002	0.0058	0.2121
L_PUT	0.3065	0.0031	0.0035	0.6167	0.3919	0.1293	0.4976	0.0194	0.0268
L_SMA	0.0055	0.2585	0.2017	0.0004	0.1047	0.0048	0.0002	0.0019	0.0176
L_THA	0.2344	0.7876	0.8087	0.1994	0.1088	0.1678	0.0011	0.0085	0.0759
R_ACC	0.3091	0.6200	0.8605	0.0154	0.2359	0.3017	0.0817	0.1963	0.9937
R_CAU	0.1352	0.5389	0.6899	0.3411	0.6633	0.0360	0.4276	0.1943	0.2750
R_CER	0.0025	0.0382	0.1035	0.0004	0.3377	0.0077	0.0014	0.0069	0.0080
R_M1	0.0011	0.0684	0.0110	0.0002	0.0289	0.0036	0.0008	0.0023	0.0126
R_PUT	0.0990	0.9790	0.0023	0.1897	0.6419	0.0112	0.3630	0.0128	0.0894
R_SMA	0.0025	0.0070	0.0085	0.0001	0.0036	0.0018	0.0000	0.0020	0.0248
R_THA	0.4070	0.9857	0.8837	0.0031	0.1063	0.1741	0.0009	0.0057	0.0581

To study the effects of medicine, the activation level of the PD group before medication is compared to that of the PD group after medication and the ROIs whose activities are significantly different are shown in Table A.2. These results clarify a controversy regarding possible M1 hyperactivation in PD [54]: it appears task-dependent, and drug dependent (Table A.2 and Fig. A.5). At lower movement speeds, the PD subjects before medication appear similar to normal subjects than those at the faster frequencies. We interpret this to be the result of a compensatory mechanism. When the PD subjects attempt the higher frequency, they may recruit a different pathway, such as the ipsilateral cerebellum (L_M1). For both the N and PB cases, the coefficient is monotonically increased as the task difficulty increases. This is consistent with our intuitive expectation. We can see that, compared with the PA case, the PB pattern is changed toward the pattern in the control group (N case), which suggests that the medicine improves the brain activity in this ROI.

Table A.2: P-value of the Activation Level across Groups - PA vs. PB

ROIS	PA vs. PB		
	F1 = 0.25Hz	F2 = 0.5Hz	F3 = 0.75Hz
L_M1	0.7989	0.0243	0.0323
R_M1	0.6926	0.3585	0.0247
R_PUT	0.0189	0.2290	0.9902

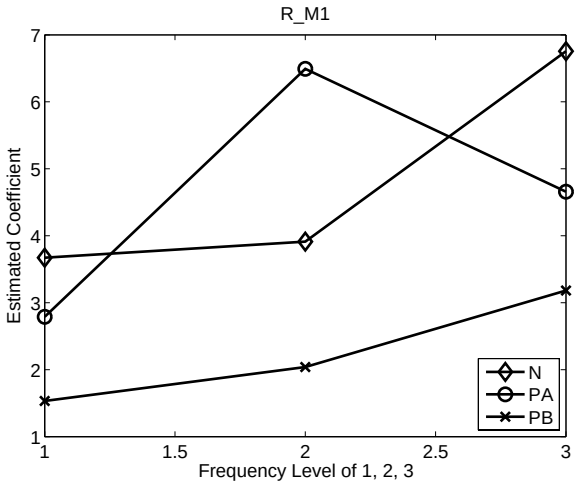


Figure A.5: The regression coefficient of the ROI R_M1 as a function of the task difficulty level.

A.5 Conclusion

This chapter proposed a spectral clustering based approach for activity detection and group comparisons in fMRI data analysis. Since no shape in the feature space assumed, spectral clustering is considered as a suitable mechanism to cluster fMRI time series. A real case study of Parkinsons disease yielded promising results, as it demonstrated high quality activation results, identifying different brain activation patterns between groups characterized by different factors such as disease and medicine.