Statistical Analysis of EEG Phase Shift Events

by

William Marshall

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Statistics

Waterloo, Ontario, Canada, 2014

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This thesis develops statistical methods for the identification, and analysis of *phase shift* events, i.e. sudden changes in the timing relationship between coupled oscillators. Phase shifts events occur in many complex systems but here the primary interest is the analysis of electroencephalogram (EEG) recordings where they have been identified as markers of information transmission in the brain; as a secondary example we analyze systems of weakly coupled Rössler attractors.

The main result, found in Chapter 2, is a novel method for estimating neural connectivity from EEG recordings based on spatio-temporal patterns of phase shift events. Phase shift events are modelled as a multivariate point process, and the ideas of Granger causality are used to motivate a directed measure of connectivity. The method is demonstrated on EEG recordings from 18 participants during three task conditions; resting, visual vigilance and auditory vigilance. Likelihood ratios are used to test the hypothesis of no Granger causal interaction between signals, and network patterns are analyzed using graph theory.

In Chapter 3 the problem of phase shift identification is formulated as a change point in the instantaneous phase. Two estimators are considered, based on the cumulative summation and the instantaneous phase derivative. Block bootstrapping techniques are used to capture the dependency structure in the signals and determine critical values for shift identification. Estimators are evaluated both on their accuracy, and temporal resolution.

Finally, detailed simulation studies are performed using realistic head models to investigate the effect of volume conduction (linear spread of electrical activity at the scalp) on phase shift analysis. Specifically, Chapter 4 investigates the effect of volume conduction on the analysis, in order to understand the limitations of the phase shift Granger causality method. Chapter 5 then investigates an approach for reducing the effect of volume conduction by using EEG source reconstruction techniques to estimate neural source activity and then identifying phase shifts with-in the brain directly from the reconstructed sources.

The primary impact is the novel method for estimating neural connectivity. Each chapter investigates a different aspect of EEG phase analysis, and together they form a complete package for estimation and interpretation of neural connectivity. Two other areas of impact are in statistical change point analysis, and behavioural psychology.

Acknowledgements

I would like to thank my supervisor Dr. Paul Marriott for his advice and support throughout this process.

I would also like to thank Dr. Sid Segalowitz, Chrissy Lackner and the rest of the researchers at the Brock University Cognitive and Affect Neuroscience Laboratory, for providing me with the EEG data for this thesis and all of their comments and discussions on this work.

Dedication

This is dedicated to my family.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

A prominent feature of brain dynamics is the transient synchronization of clusters of neurons into *neural assemblies* [35, 138, 119, 261, 280]. Such behaviour suggests a temporal coding for information transmission through the brain [286, 232]. Strongly synchronized systems have relatively trivial dynamics [19, 235], and could not account for the range of possible brain functions [277]; in fact, excessive synchrony has been associated with pathological conditions such as schizophrenia [176] and epilepsy [175, 190]. Thus the processing of information requires not only synchronization, but also the ability to spontaneously desynchronize clusters of neurons [94, 93]. These spontaneous desynchronizations, or *phase shifts*, may contain valuable information about the functional nature of neural synchronization [277, 273, 167].

One way to study phase dynamics is to use a high temporal resolution brain imaging technique such as electroencephalogram (EEG) recording [200]. The use of EEG (sampling rate $\sim$ 1kHz) recordings is well suited for phase shift analysis because it has the required temporal resolution to identify the transient events, unlike other techniques such as fMRI (function magnetic resonance imaging, sampling rate $\sim$ 1Hz). The unifying theme of this thesis is the development of techniques for the statistical analysis of phase shift events from EEG recordings.

A recently introduced method of phase shift analysis from the EEG literature is *phase reset* [277, 273]. In phase reset analysis, there is a spontaneous desynchronization (phase shift) which follows a period of prolonged synchrony (phase lock); signals are partitioned into periods of phase locking and phase shifting behaviour, after each phase shift the signal enters a new period of phase locking. In collaborative work with the Brock University Cognitive and Affective Neuroscience Lab, we have performed a behavioural study, investigating the relationship between phase reset metrics with composite measures of anxiety and aggression in an adolescent population. Results showed significant correlation between both aggression and anxiety traits and the phase reset metric, even after controlling for the results of a traditional coherence analysis. These results have been accepted for publication in *Brain and Cognition* [167].

The primary contribution of this thesis is a novel measure for assessing connectivity between EEG signals based on spatio-temporal patterns of phase shift events. Identified phase shift events are modelled as a multivariate point process [59], and the framework of *Granger causality* [116, 241, 15, 160] is used to estimate a *directed* measure of connectivity between signals, which we have termed phase shift Granger causality (PSGC). The proposed measure is quite different from standard methods in EEG connectivity analysis [205, 190, 245, 201, 264]; instead of quantifying the average strength of connectivity, our measure investigates patterns in discrete desynchronization events. There are a number of benefits to the PSGC measure over a traditional coherence metric [205]; it identifies non-linear relationships which may not be picked up by traditional coherence metrics, it works well in the presence of noise, it makes use of the full temporal resolution of the EEG recording, it allows for statistical validation of connectivity networks from an individual recording.

After initial exploratory work on developing the PSGC measure, feedback received from experts in the field focused on two main aspects of the analysis; the identification of phase shift events, and the effect of *volume conduction* (VC) on estimated connectivity (see §1.5.1). For the former, concerns exist over the subjective selection of a threshold parameter, and hence the repeatability of the study. With VC, there are questions as to

whether the results of PSGC analysis could be interpreted as true connectivity between neural regions, or if it is confounded by electric field spread across the scalp.

Based on the above feedback, the next contribution of this thesis is an investigation into methods of phase shift identification from observed time series. The methods currently used in the EEG literature [37, 277, 167] involve an unwanted subjective aspect, so there is a need for a reproducible, objective method. We formalize the process of shift identification as a change-point in the instantaneous phase variable, and compare the phase derivative methods used in the EEG literature to a more traditional cumulative summation (CUSUM) based method from the change-point literature [179]. In addition to the application to EEG phase shift analysis, these results also have application in other fields which consider complex systems of oscillators, which occur, for example, in biology [187, 20], interferometry [85, 123] and geology [269, 109].

The thesis then considers the effect of VC on PSGC analysis. Essentially, the question which must be answered is, if signals are coupled in the desired way, but then linearly mixed using a biologically plausible lead field matrix, what is the result of connectivity analysis applied to the mixed signals? A computationally intensive simulation study is designed as a 'proof-of-concept' for the PSGC measure under conditions similar to those in EEG recordings, including linear mixing and high levels of noise. Based on the results of the study, it is determined that the existence of spatio-temporal 'biological' noise helps to attenuate the spread of shift events due to VC, and so the PSGC measure is well suited for application in low signal-to-noise ratio (SNR) situations, which are ubiquitous in EEG recordings [58, 72]. The information from the simulation study has been packaged together with the results from the original application of the PSGC measure.

One potential way to compensate for the effect of VC in EEG connectivity analysis is to apply source reconstruction techniques, and then estimating connectivity directly from the estimated neural sources. This process also increases the spatial resolution of the analysis (a weakness of EEG [200]). As an initial step in this process, the thesis investigates the joint identification and localization of phase shift events in neural space using source reconstruction techniques. There are many options for solving the underdetermined EEG

inverse problem [246, 193, 215, 70, 288, 191, 131, 136, 72, 18, 62, 282], in this work we focus on identifying the reconstruction technique which is most suitable for application in phase shift identification. This work is currently in preparation to be submitted to the *Journal of Computational Neuroscience*, it has implication for both the analysis of EEG recordings using PSGC, but also other EEG analyses based on phase shift events, such as the phase reset analysis.

In Chapter 2, I introduce the novel technique for estimating connectivity from EEG recordings, called Phase Shift Granger Causality. Identified phase shift events are modelled as a multivariate point process, and the framework of Granger causality is applied to produce a directed measure of connectivity between signals. The PSGC analysis provides additional insight into the nature of phase synchronization, and its role in information transmission in the brain, by exploring spatial and temporal patterns in spontaneous phase desynchronization. To demonstrate the proposed measure, PSGC analysis is applied to 18 recordings from an adolescent population during three conditions, resting, auditory vigilance and visual vigilance.

Chapter 3 is dedicated to a statistical treatment of phase shift identification. Current applications often employ a threshold technique for shift identification; however, since there is no available information to validate the threshold, selection is often done subjectively. I formalize this problem as a change-point problem for a time-dependent instantaneous phase variable. Two potential estimators are discussed for shift identification, with special emphasis on both the accuracy and temporal resolution of each estimator.

In Chapter 4 I focus on the instantaneous spread of electrical activity across the scalp, a process known as volume conduction (VC). The effects of VC are studied for both phase shift identification and the localization of neural connectivity. Simulation using realistic head models is used to assess the spread of shift identification due to VC.

In Chapter 5, I compensate for the effects of VC by inverting the scalp measurements to

recover the original sources within the brain, this process is known as source reconstruction (see Section 1.6). The focus is on a Bayesian model for source reconstruction, and the ability for such a model to accurately localize phase shift events. Finally, in Chapter 6 I give some concluding remarks on the use of phase shift events and outline some outstanding questions and future work.

I begin the first chapter with history of EEG, early use of the technique and important results which helped to shape the field. I then explore the effect of the digital era on EEG recordings, including the benefits and new problems which arise from the use of computers, the chapter concludes with a discussion of new applications at the forefront of EEG research.

1.1 Data Primer

This thesis is primarily concerned with the *phase shift* events which separate periods of synchronization in weakly coupled oscillators. In this first section, I introduce the notion of *instantaneous phase* which is necessary to accurately estimate and localize such events, then explore and visualize the instantaneous phase of several sources of data used throughout the document.

1.1.1 Instantaneous Phase

Phase can be considered as the value which defines the initial state of an oscillator, and many linear techniques, such as the Fourier transform, assume a constant phase value over the window of estimation [28]. Nonlinear methods are used to estimate time-dependent measures of *instantaneous phase* which capture moment-to-moment changes in the state of an oscillator. Instantaneous phase can be defined using either the Hilbert [101] or Wavelet transformations, both have been shown to produce similar results [227], here we use the Hilbert definition.

The Hilbert transform can be used to define the analytic extension of the signal, eliminating power in the negative frequency bands. With all the power contained in the positive frequencies, the direction of oscillation is known and the phase can be unambiguously estimated. For a continuous time real valued process x_t, the analytic extension is defined as [101]

$$\chi_t = x_t + i\tilde{x}_t,$$

where $\tilde{x}_t$ is the Hilbert transform of the signal. The Hilbert transform of x_t is defined to be

$$\tilde{x}_t = \frac{1}{\pi} PV \int_{-\infty}^{\infty} \frac{x_\tau}{t-\tau} d\tau,$$

where the integral is taken using the Cauchy principal value due to the potential singularity at $t = \tau$. The Hilbert transform is a convolution of the signal x_t with the function $\frac{1}{\pi t}$. The power spectral density (PSD, see §1.4.2) of $\tilde{x}_t$ is equal to the product of the PSD of x_t and the PSD of $\frac{1}{\pi t}$,

$$\mathbb{F}\left(\frac{1}{t\pi}\right) = -isgn(\omega) = \begin{cases} e^{i\pi/2} & if \quad \omega < 0, \\ 0 & if \quad \omega = 0, \\ e^{-i\pi/2} & if \quad \omega > 0. \end{cases}$$

Thus the Hilbert transform is simply a $\frac{\pi}{2}$ shift in the phase of the original signal.

Working with the analytic extension of the signal does not change the PSD; however, additional information can be recovered using the tools of complex analysis. The instantaneous amplitude and phase are defined as the magnitude and argument of the analytic extension. For example if χ_t is the analytic extension of the real valued signal x_t, then the instantaneous amplitude A_t and phase ϕ_t are defined such that

$$\chi_t = A_t e^{i\phi_t}. \tag{1.1}$$

Conditions are given for a signal to have a physically meaningful interpretation of the instantaneous phase and amplitude in [48]. Here we require that the signal be described by a small spectral band centred on the main frequency component, and that it satisfies

the *narrowband* approximation, that is, it is one whose relative rate of change in amplitude is small compared to rate of change in phase.

In practice, the complex demodulation algorithm is used to estimate the instantaneous phase of an observed signal [28, 113] (see Appendix G for additional details).

Definition 1.1. *For a given δHz low-pass filter $\mathbb{H}[\cdot]$, the instantaneous phase of an observed signal x_t in the frequency band $(\omega - \delta, \omega + \delta)$ is,*

$$\hat{\phi}_t = \tan^{-1}\left(\frac{\mathbb{H}[x_t \sin\{\omega t\}]}{\mathbb{H}[x_t \cos\{\omega t\}]}\right). \tag{1.2}$$

The δHz low-pass filter isolates the frequency band of interest, this ensures that the calculated instantaneous phase can be meaningfully interpreted as the phase of the $(\omega-\delta, \omega+\delta)$ component of the signal [48]. Also note that any discontinuities due to the inverse tangent function are removed using *phase straightening*, by searching for discontinuities in the instantaneous phase and, when one is found, adding or subtracting 2π based on the direction of the discontinuity.

As a simple example of estimating instantaneous phase using the complex demodulation algorithm, consider a noisy oscillator with a change point in amplitude and phase,

$$x_t = \begin{cases} \sin\left(2\pi f_0 t/T\right) + \epsilon_t, & t < 3T, \\ 2\sin\left(2\pi f_0 t/T\right) + \epsilon_t & 3T \leq t < 6T, \\ 2\sin\left(2\pi f_0 t/T + \pi/3\right) + \epsilon_t & 6T \leq t, \end{cases}$$

where $T = 32$ is the sampling rate, $f_0 = 4$ the frequency of the oscillator, and ϵ_t are independent $G(0, 0.5)$ random variables. The raw data and estimated PSD for this example are shown in Figure 1.1. The change-point in amplitude can be clearly identified from the raw data at $t = 3T$, but the effect of the phase shift is much more subtle. The PSD is able to correctly identify that the primary component is at 4Hz, but there is no suggestion of a phase shift event. In Figure 1.2, we can see that the change points in both phase and amplitude are captured by the complex demodulation algorithm, though there is some

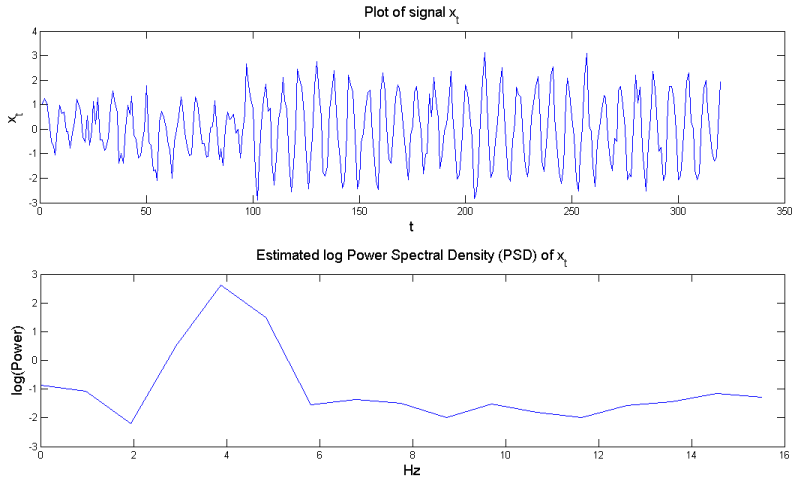

Figure 1.1: Raw signal of the noisy oscillator x_t (top), and estimated PSD using Welch's method [296] (bottom). From the plot, the change in amplitude is clearly observed at $t = 3T$, but the change in phase at $t = 6T$ is much more subtle. The existence of a 4Hz component can be clearly identified from the PSD, but there is no information about the phase.

residual oscillatory effect in the estimates (this effect is further explored in §3.1.1).

An important aspect of the instantaneous phase analysis is that the conditions be met for an interpretation of instantaneous phase; an example of this potential problem is given in Appendix G.

1.1.2 Rössler Attractor

The Rössler attractor is a simple to describe mathematical system which displays relatively complex dynamics. Systems of weakly coupled Rössler attractors are primarily phase locked, but are prone to spontaneous desynchronizations [235, 222]. Simulated data from systems of weakly coupled attractors provide an additional application to validate

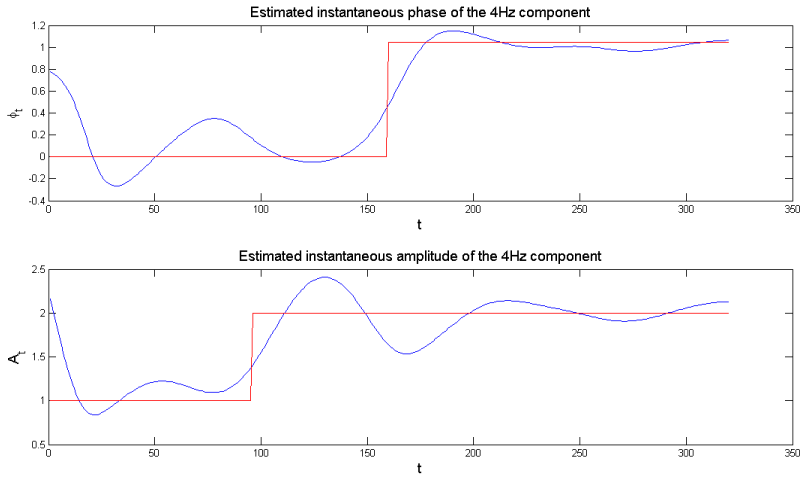

Figure 1.2: Estimated instantaneous phase (top) and amplitude of the 4Hz component estimated using the complex demodulation algorithm.

the methods in this thesis for which the truth shift events can be ascertained.

The Rössler system is a set of three ordinary differential equations [236], with parameters a, b and c, which is linear except for a single bi-linear term. For certain values of the parameters, the attractor primarily rotates around the origin in the x-y axis, while showing spontaneous bursts in the z-direction. Sample trajectories generated with a=b=0.2 and c=5.7 [236] and random initial conditions are shown in Figure 1.3; a burn-in period of 30 seconds is thrown away to remove transient activity.

Here we explore the two coupled Rössler attractors,

$$\begin{aligned} \dot{x}_i &= \omega_i(-y_i - z_i) + C(x_i - x_{|i-1|}), \\ \dot{y}_i &= \omega_i(x_i + a y_i), \\ \dot{z}_i &= \omega_i(b + z_i(x_i - c)), \end{aligned}$$

for $i = 0, 1$, where $\dot{x}$ represents the temporal derivative of x, ω_i the attractor frequency and C is the coupling strength.

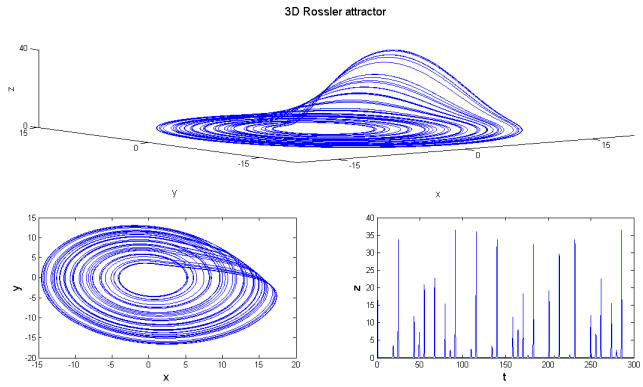

Figure 1.3: Sample trajectory of a Rössler attractors with $a = b = 0.2$, $c = 5.7$.

Trajectories primarily oscillate in the x-y plane, so there is a natural definition for the phase of this attractor using Poincare maps [235], which will be take as the true phase of the system,

$$\phi_t = tan^{-1}\left(\frac{y_t}{x_t}\right). \tag{1.3}$$

Instantaneous phase variables can also be defined in the Hilbert transform sense, using observables of the system. The complex demodulation algorithm is applied with a centre frequency of $\omega = 9.25$ Hz (see Figure 3.2) and a bandwidth of $\Delta\omega = 0.15$ Hz; resulting phase differences between the two coupled oscillators are shown in Figure 1.4, for both the true phase and the Hilbert phase with observables $h(x_t, y_t, z_t) = x_t$ and $h(x_t, y_t, z_t) = y_t$.

1.1.3 EEG Data

In order to assess the proposed methods in this thesis, several EEG recordings have been provided by Dr. Segalowitz and the *Brock University Cognitive and Affective Neuroscience Lab*. Recordings were obtained from 18 participants (9 male, 9 female, aged 12-14 years) during resting, auditory and visual tasks. In the first task, participants rested for four minutes while fixating on a cross at the middle of a black computer monitor in a dimly

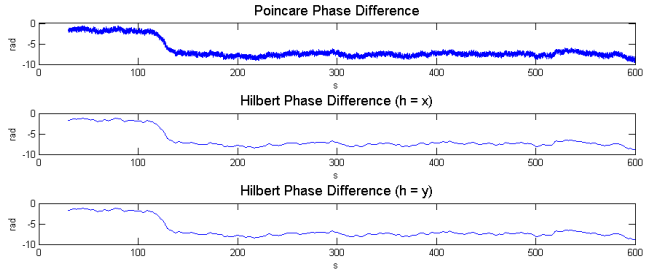

Figure 1.4: Coupled Rössler attractors with $a = 0.15$, $b = 0.2$, $c = 10$, $C = 0.12$, $w = 2\pi 9$, $dw = 0.0675$. Phase difference plots for three definitions of phase; top: Poincare phase, middle: Hilbert transform of x_t, bottom: Hilbert phase of y_t. In all three definitions, there is a clear locking / shifting dynamic, with a phase shift at approximately $s = 130$ seconds.

lit room. The visual attention task was a traditional Erikson Flanker task [83]. The task was to discriminate two stimuli by pressing the corresponding button for each. Stimuli were presented for 200 ms followed by a variable intertrial interval (ITI) of 800 to 900 ms. In a selective auditory attention task, two 200 ms digitized tones were presented to one ear at a time in a random order: a 1000 Hz standard tone (88% probability) or a 2000 Hz deviant tone (12% probability). Sounds were presented with a variable ITI of 600 to 800 ms. Participants were instructed to attend to one ear only and to ignore all sounds presented to the other ear and to respond by pressing a key when they heard the deviant tone in the ear of attention, and not to respond otherwise. They visually fixated on a cross at the centre of the computer screen during the task. The visual and auditory vigilance tasks took approximately 15 minutes each to complete.

EEG recordings were obtained from 121 scalp sites at a sampling rate of 500 Hz. The recordings were reduced to a set of 19 standard sites (see Figure 1.7) representing regions of the left hemisphere (Fp1, F7, F3, C3, P3, T3, T5, O1), the right hemisphere (Fp2, F8, F4, C4, P4, T4, T6, O2) and midline (Fz, Cz, Pz). Impedances were maintained below 30 kΩ throughout recording. Data were re-referenced offline to the average of all sites, filtered offline (1-30 Hz) and corrected for eye movements using the Gratton and Coles

procedure [118]. An automated artifact rejection procedure was used in addition to manual examinations of the data.

Figure 1.5 shows 10 second sample waveforms from two electrodes in a single EEG recording during the visual task. The corresponding PSDs (estimated from the entire duration of the signal) are also shown, we see that there is substantial power across many frequency bands, and also the effect of the 30Hz low-pass filter applied during pre-processing stage.

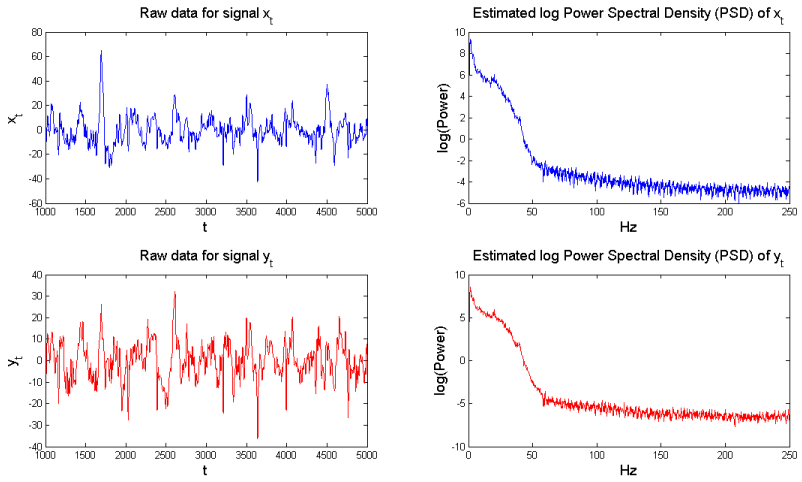

Figure 1.5: A 10 second sample from two electrodes during the visual vigilance task, and corresponding estimated PSDs.

In Figure 1.6 we see the estimated instantaneous phase of (8-10)Hz frequency component from each signal, as well as the instantaneous phase difference. In these plots we see the distinct phase locking and then phase shifting dynamic which the *phase reset* measure attempts to capture. It can also be seen that some shifts are emphasized in the instantaneous phase difference while others are attenuated.

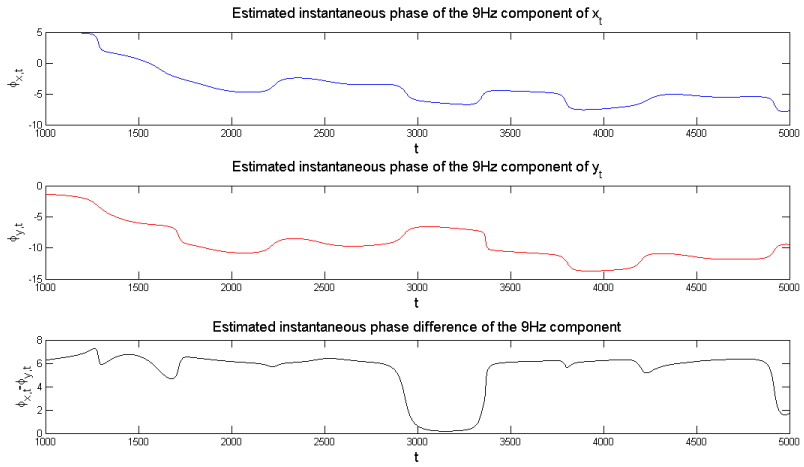

Figure 1.6: Estimated instantaneous phase of x_t (top) and y_t (middle). Both signals show prolonged periods of constant phase values (locking) and spontaneous jumps in the phase (shifts). The bottom panel shows the instantaneous phase difference; in the phase difference some of the shift events are exaggerated, while others are attenuated.

1.2 History of EEG

Electroencephalogram recording is a brain imaging technique which was first used on humans by Hans Berger in 1929 [25]. It is a non-invasive technique which measures electrical activity in the brain with sensors (electrodes) placed against the scalp. The number and placement of electrodes differ depending on the application, however in 1958 the 10-20 electrode placement system [147] was developed to help standardize reported results (see Figure 1.7). Modern electrode nets record many more channels than the 19 suggested by the 10-20 system (see Figure 1.8), though results are still often reported using 10-20 terminology.

In 1936 the Massachusetts General Hospital [117] was the first hospital to begin using EEG recordings for clinical purposes. Recordings were used to diagnose localized brain disorders, for example epilepsy [107] or brain tumors [27]. New imaging techniques such

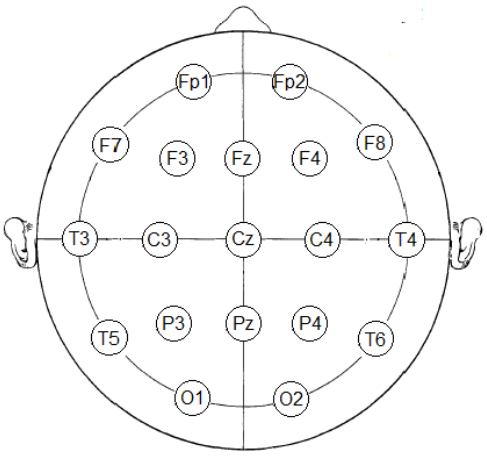

Figure 1.7: Visualization of the international 10-20 electrode placement system. The outer electrodes are spaced in 10% and the inter-electrode distance is 20% of the nasion to inion distance of the head (for a row with 5 electrodes, they are placed at 10%-30%-50%-70%-90%). The first digit in the name identifies the region of the cortex (frontal, central, temporal, parietal, occipital) and the second digit describes the hemisphere (odd on the right, even on the left, z on the midline).

as fMRI or x-ray computed tomography have since replaced EEG for many localization applications due to superior spatial resolution. However, using EEG recordings for scientific, rather than clinical, purposes continues to be popular because it has greater temporal resolution than other imaging techniques, as well as being non-invasive, cost effective and portable.

In the following, I will review the most important physiological and technical details which underlie EEG recording. I will give a historical review of the different methods have EEG analysis which have been used, and how they have evolved with changing technology. I will then introduce two modern problems in EEG analysis (source reconstruction and measuring connectivity) which will be investigated in later chapters. Finally I will give a

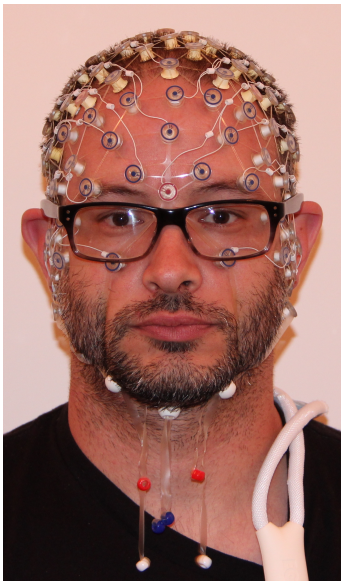

Figure 1.8: A modern EEG cap with 128 electrodes. Electrodes come as far down the head as reasonable to try and pick up the activity originating from the bottom of the brain. There are electrodes directly above and below each eye, which can be used to isolate activity due to eye movements.

brief preview of what the future holds for EEG analysis.

1.3 Electricity in the Brain - Basis for EEG

Information is transferred through the nervous system by billions of cells called *neurons* or nerve cells. A typical neuron has a main cell body surrounded by tiny branch-like fibers called *dendrites*, which are used to receive information, and a longer *axon*, fiber which branches out at the end to communicate information to other neurons. Figure 1.9 is a diagram depicting a typical neuron. When a neuron is stimulated by pressure, heat or light it may discharge a brief electrical current, which travels through its axon, called an *action potential* or spike. Once the action potential reaches the ends of the axon branches it travels across a small gap called a *synapse*, and is received by the dendrites of other neurons.

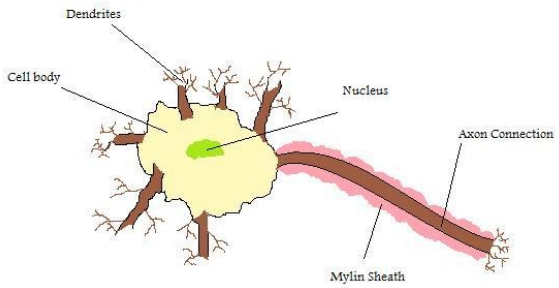

Figure 1.9: A diagram of a typical neuron. The neuron consists of a main cell body, many bushy dendrites (input) and a long axon channel with insulating sheath (output).

Electrical recordings from within the cortex show that the extra-cellular current which occurs as the signal crosses the synapse has the largest magnitude of all electrical brain activity (which can be up to 30 mV) [5]. Taking into account the distance to the electrode and the conductivity of brain tissue, many signals likely do not have a measurable effect at the scalp; however, signals such as the auditory potential, which originate deep in the brain stem are able to produce measurable effects at the scalp by traveling through highly conductive spinal fluid [150]. Evidence suggests that most electrical activity at the scalp is due to clusters of *pyramidal neurons*, called *cortical columns* due to their columnar structure and axons which project perpendicular to the surface of the cortex [81, 173]. Moreover, signals from these columns can be amplified by synchronous activation patterns over periods of time [43]. These transient groups of functionally connected neurons are now called *neural assemblies* [138].

1.4 EEG Analysis

The goal of EEG analysis is to learn about brain function. This is accomplished by using information from theoretical models, clinical observations and spatio-temporal EEG

patterns [57]. This document will focus on using statistical techniques to identify spatio-temporal patterns in EEG recordings.

A sample plot of an EEG recording is shown in Figure 1.10, with signals recorded from five locations (Fp1, F3, C3, P3, O1). The recording is taken from a person at rest, while fixated on a star pattern screen saver. The signal from location Fp1 which is closest to the eyes, shows distinct positive spiking waveforms (one at $0.5s$, $2.5s$ and two at $5s$) which are characteristic of blinking and provide no information about brain activity. As the location of electrodes moves away from the eyes towards the back of the head, the blink activity is attenuated and blurred (the two blinks at 5s appear as one) and once the electrodes move below the eyes (O1) the effect is negative. The sub panel of the figure shows a typical second of EEG data, without any blinking artefacts. Muscular artefacts such as those due to blinking are just one of many sources of noise in EEG recordings, a more detailed description of noise in EEG recordings is given in Appendix B.

1.4.1 Early analysis

Early EEG recording systems used a string galvanometer [80] and later vacuum tubes [90] to measure and amplify electrical potentials at the scalp, the resulting analogue signal was attached to a pen motor which drew the recording onto paper. Systems which would draw the EEG signal onto magnetic strips were eventually developed, allowing the recording to be played back [44].

Due to the nature of EEG recordings at this time, analysis was done visually and focused on the qualitative features of the signal. Features of the EEG signal were classified as either paroxysmal (transient, bursting) or on-going (background, spontaneous) activity. A simple feature of EEG recordings which was used is the dominant frequency, measured by counting the number of oscillations in one second. These measurements were not exact, so EEG activity was described in terms of several predetermined frequency bands; delta (0-4Hz), theta (4-8Hz), alpha (8-13Hz) and beta (13+Hz). It was difficult to quantify the amplitude of frequency component, since the peak-to-peak amplitude of the signals

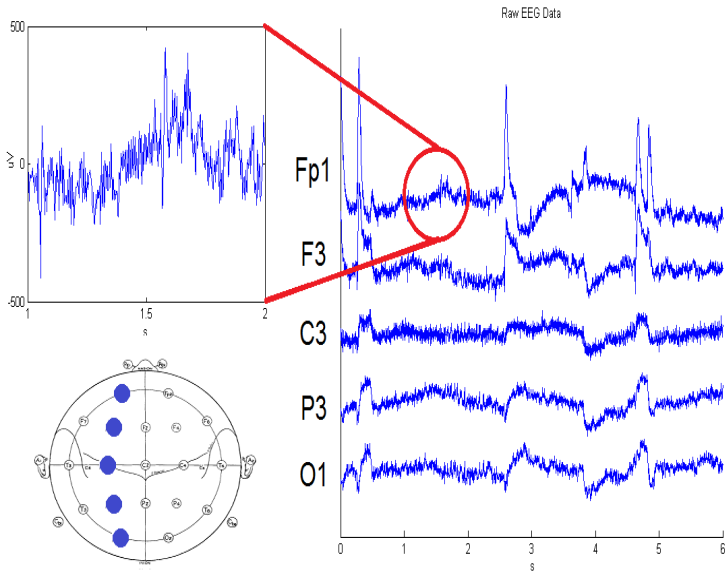

Figure 1.10: The main figure shows six seconds of EEG from a resting task, where the participant fixated on a star pattern screen saver (data provided by Brock University Laboratory of Cognitive and Affective Neuroscience). The recordings are from five locations arranged from the front to the back of the head. In the Fp1 recording there are many sharp positive spikes, these waveforms are characteristic of an eye blink. As the signal propagates across the skull, the effect of the eye blink is attenuated, and at the back it has shifted to negative due to being physically below the source.

changed over time due to noise or other components in the signal. One approximation used was to create an envelope for the signal by joining together all the peaks/troughs and taking an average of the resulting lines [291]. Another measure of interest was the fraction of time that a recording would spend oscillating in a specific frequency band [110].

There are also many more complicated waveforms that commonly exist in EEG recordings, a few examples will be described here but a more detailed treatment can be found in [47]. The alpha rhythm is a waveform with frequency in the alpha band, it was the first EEG waveform reported by Hans Berger [25] and has also been called the Berger waveform. The alpha rhythm is most prominent in the occipital region of the cortex and is

strongest during eyes closed relaxation, the rhythm is attenuated during attention (especially visual) [47]. The K-complex is a transient waveform that occurs during sleep, either spontaneously or in response to stimuli. The waveform has a sudden jump of around 100μV, then a slow return to baseline followed by several seconds of activity at approximately 14Hz [65, 237].

Topographic maps of the scalp potentials were used to visually identify spatial patterns in EEG, however these maps were cumbersome to create from analogue recordings. One method for making maps is to use bipolar recordings (the difference between two linked electrodes) directly to measure the potential gradient in one direction on the scalp. If many electrodes were arranged in a line, bipolar recordings could be used to localize the maximum of the potential by identifying the point where the gradient shifted from positive to negative [3].

1.4.2 Digital Recording system

Often visual inspection was not sufficient to discriminate between components [68]; furthermore, more complex analyses were very cumbersome with analogue recordings. To create digital EEG recordings, analogue-to-digital converters are used to sample potential values at regular time intervals and then assign them to a discrete set of possible voltage values. Digital EEG recordings were in development for some time, it was not until the 1970s that they overtook analogue as the recording of choice [106].

Digital methods opened new possibilities for analyzing the spectral content of EEG waveforms. Based on the ideas of Fourier analysis, the fast Fourier transform (FFT) [56] is an efficient algorithm which is used to estimate the *power spectral density* (PSD) of a signal. The PSD quantifies both the power and phase for each frequency component of the signal in an objective way.

Definition 1.2. *The power spectral density* (PSD) of a stationary process X_t is

$$h_X(\omega) = \sum_{s=-\infty}^{\infty} \delta_s e^{-is\omega}, \quad |\omega| < \pi,$$

where δ_s is the auto-covariance function

$$\delta_s = Cov(X_t, X_{t-s}), \quad s \in \mathbb{N}.$$

Creating scalp topographies was made more practical with digital technologies. Computer methods used interpolation to create detailed contour plots of the scalp potential which are useful for localizing specific waveforms. However, ambiguities can arise in the topography when interpolating scalp potentials [29], due to projecting a 3D volume onto a 2D surface. A similar spatio-temporal method was developed by [230] called chronotopograms, where the contour plot was created with one spatial dimension and one temporal dimension. This gave an idea of the topography of a single line of electrodes as it changed over time.

Event Related Potentials

Event related potentials (ERPs, previously known as evoked potentials) are short EEG recordings which are time locked to an event, usually the onset of a stimulus. The trials for an ERP are repeated many times and then averaged, maintaining the waveform of interest while reducing the effect of unrelated background noise. The ERP was first described by [67] using superimposed images to emphasize the signal, however with digital recordings a simple average is used to generate the waveform.

The ERP for basic sensory stimuli have been identified (visual [128], auditory [150], olfactory [163], somatosensory [292]) and are well studied. The visual evoked potential (VEP) is a pattern evoked by changing visual stimuli, such as a checkerboard pattern with alternating black and white squares. The primary component of a VEP is a large positivity (up to 30μV) with a random latency which occurs on average 100ms after the

stimulus and it is strongest at the Oz region of the scalp. Features of the VEP such as its latency and amplitude have been shown to dependent on the intensity of the stimulus [127].

The are certain components of ERP waveforms which are pervasive across many different stimuli. The N1 (or N100) is a negative deflection in the ERP that occurs approximately 100ms after the onset of an unexpected stimulus, which is localized in the Fz frontal regions (see [197] for a review). The amplitude of the N1 depends on the expectation of the stimulus, for regularly occurring stimuli the amplitude is reduced (called habituation) and it is altogether absent when the participant is given control of stimulus onset [66].

In events where it is required for the participant to make a decision regarding the observed (not necessarily visual) stimulus, the P3 (or P300) is a large positive deflection in the ERP that occurs roughly 300-350ms after the stimulus [267], and it is maximal on the mid line of the central/parietal regions (Cz, Pz). A popular visual task for evoking the P300 is the Eriksen flanker task [83], where the participant must quickly identify the centre letter in a group while it is flanked with similar or dissimilar letters (see Figure 1.11).

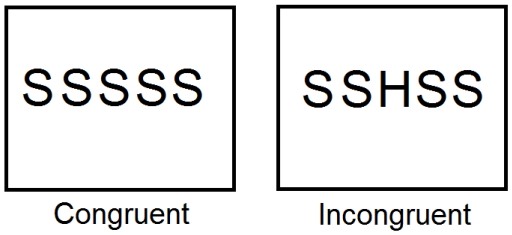

Figure 1.11: The visual stimuli for an Eriksen flanker task [83]. In the congruent task, the target letter is the same as the ones flanking it, whereas the incongruent task has a different target letter. In repeated trials, this incongruence causes an increased rate of errors, allowing researchers to analyze the waveform associated with error detection.

Coherence

The PSD and scalp topographies are used to learn about functional segregation, the idea that distinct regions of the brain which are responsible for specific tasks. It is known that functional segregation is a key aspect of brain function in mammals [86, 177, 301]; however, it is not the only aspect of brain function [169]. *Functional integration* is another important feature of brain function, it is the coordination of multiple functionally segregated brain regions to achieve a unified response [88, 119, 261].

To measure functional integration, measures were introduced to quantify *functional connectivity*, a statistical dependence between EEG signals from different areas of the scalp. The first prominent measure of functional connectivity was *coherence*, a bivariate spectral quantity which measures the 'correlation' between the signals at specific frequencies [258, 290].

Definition 1.3. *The cross spectral density* (CSD) of a bivariate stationary process (X_t,Y_t) is

$$h_{XY}(\omega) = \sum_{s=-\infty}^{\infty} \delta_s e^{-is\omega}, \quad |\omega| < \pi,$$

where δ_s is the cross-covariance function

$$\delta_s = Cov(X_t, Y_{t-s}), \quad s \in \mathbb{N}.$$

Definition 1.4. *The coherence* between variables in a bivariate stationary process (X_t,Y_t) is

$$\gamma^2_{XY}(\omega) = \frac{|h_{XY}(\omega)|^2}{h_X(\omega)h_Y(\omega)}, \quad |\omega| < \pi.$$

In practice Welchs' periodogram is used to estimate coherence [296], which uses the FFT to calculate P/CSD for many segments of data and then averages the P/CSDs (such as in ERP studies). Using the average reduces the variance of the estimator, which is important for noisy EEG recordings, however the drawback is a lower frequency band resolution in the solution. The papers [205, 204], are a comprehensive study of the properties of coherence as well as how it is affected by the noise in EEGs; main results were the

existence of spurious high coherence estimates between scalp recordings (low-moderate interelectrode distance $< 12cm$) when compared to cortical recordings, due to the effect of VC across the scalp.

There exists structural subcortical connections between distinct regions of the cortex [34] called intracortical (or cortico-cortical) connections. In [274], Thatcher uses coherence to map functional development of these intracortical connections from birth to 16 years. The results showed a cyclic pattern between growth and reorganization in cortical development.

Although coherence is a successful measure of functional integration, it has several limitations.

- Stationarity: Existence of a PSD depends on the stationarity of a process. There is evidence that EEG recordings longer than 12 seconds are non-stationary [55]
- Linearity: Coherence is based on linear correlation in the process, if information is being transferred in a nonlinear manner, it may not be identified.
- Direction: EEG coherence does not identify which signal (if any) is driving the flow of information.

1.5 Modern Analysis - Measuring Connectivity

A current scientific model for how functional integration is achieved in the brain is based on the concept of *neural assemblies*. Neural assemblies are transient dynamic groups of functionally connected neurons. Neurons within an assembly are able to interact more strongly with each other via an amplified feedback mechanism called *reentry* [261, 120]. The process of reentry requires specific timing relations (frequency depends on transmission times between specific neurons), which requires neurons within an assembly to synchronize their firing rates. This has led to increased use of phase variables, specifically their

synchronization, to measure connectivity in EEG recordings. The use of phase variables to measure connectivity will be revisited in Chapters 3 and 4.

1.5.1 Volume Conduction

The goal of connectivity analysis is to estimate connectivity between neural regions from EEG recordings, an implicit assumption is that the EEG recording at an electrode is due to the neural regions directly below the electrode. In fact, the electric field due to neural activity projects across the entire surface of the head, an effect known as *volume conduction* (VC).

At any point in time, the electric potential on the surface of a volume is a linear combination of the potentials due to each generator within the volume. This field spread of electrical activity is governed by Maxwells equations of electricity and magnetism (see Appendix A, solving for the potential due to a given source configuration is referred to as the forward problem (see §1.6.1). From the solution to the forward problem (Equation A.2) we see that the contribution to the electric field are instantaneous (due to the quasi-static approximation [129]), and decrease quadratically with distance from the source.

Despite the above concerns, it may still be reasonable to interpret the EEG activity as being primarily due to activity directly under the electrode (see §1.3). Cortical columns are clusters of similarly oriented neurons at the cortex (shortest distance from the scalp), and may fire synchronously if recruited into a *neural assemblies*. It is reasonable that activity at the scalp is primarily due to this perpendicularly oriented synchronous activity, although certain activity does spread further (for example, the auditory evoked potential reaches the scalp via conductive cerebral spinal fluid pathways [150]).

The difference between estimated connectivity directly from neural sources and the corresponding estimates from scalp recordings is an effect of VC; this can be thought of study

error, a systematic bias in connectivity estimates due to recording from the scalp instead of from cortical activity directly. Signals recorded at the scalp show increased correlation compared to the underlying sources [205]. Furthermore, the effects of VC are modulated by noise [249], and so with-in subject differences may be a result of changing noise conditions rather than changes in underlying connectivity.

Modern measures of connectivity look to overcome the inherent limitations of EEG coherence, as well as reduce the effect of VC on estimated connectivity. There have many proposed methods used for inferring spatio-temporal patterns in EEG signals (for a review see [63, 226]), the most popular ones will be discussed here.

1.5.2 Directed Coherence

Granger causality is a statistical definition of causality originally described by Granger in 1969 [116], which is used in time series analysis. Conceptually, Granger causality is about predictability and more specifically whether knowledge of the past of one signal can improve predictors for another signal. Based on the idea that the future cannot cause the past, there is a natural direction of information flow when measuring Granger causality.

Definition 1.5. *Let X_t and Y_t be two jointly stationary stochastic processes, $P(X_t)$ the optimal predictor of X_t and U_t be all other information at time t. It is said that Y_t Granger causes X_t ($Y_t \Rightarrow X_t$) if knowledge of Y_t causes a reduction in the prediction error of X_t after all other information has been taken into account,*

$$Var(P(X_t \,|\, \mathbb{X}_t, \mathbb{Y}_t, \mathbb{U}_t) - X_t) < Var(P(X_t \,|\, \mathbb{X}_t, \mathbb{U}_t) - X_t).$$

where the font type $\mathbb{X}_t$ represents the history of the process,

$$\mathbb{X}_t = \{X_{t-1}, X_{t-2}, ..., X_0\}.$$

The above definition is a general framework and it does not provide specific details for measuring or testing for Granger causality. The optimal predictor $P()$ will depend on the

model used to describe the processes, often it is taken to be the optimal linear predictor of X_t. There is also U_t which is defined as all other information at time t, its purpose is to eliminate possible confounding variables and so ideally U_t will contain as much information as possible. What variables are included in U_t is determined by both the availability of information on potentially confounding variables and the ability of the model to compensate for different kinds of information. A true causal relationship can only be inferred when all possible confounders have been taken into account; however, even if not all confounders are available, the Granger causality analysis can still be used both for prediction and to identify directed measures of association. Two processes are said to have a feedback relationship if both $X \to Y$ and $Y \to X$. In the case of a feedback relationship, Granger causality does not provide any information about the origin of information in the system.

In order to determine the direction of information flow between EEG signals, the ideas of *Granger causality* can be used In Granger's original paper, he introduced a statistical test for causality between two processes based on a bivariate auto-regressive model [116], the measure has a convenient representation as a decomposition of the traditional coherence function and was later formalized as *directed coherence* (DC, see Appendix E). Analysis of EEG recordings using DC was first applied over a decade after Granger's original paper [241].

There are also extensions of DC such as the directed transfer function [154] which allows causality to be tested in a full multivariate system and *partial directed coherence* (PDC) [245] which accounts for the effect of other signals. To address the issue of stationarity, adaptive time varying version of both DC [140] and PDC [9] have been applied to EEG signals.

Calculation of measures such as DC and PDC is based on fitting a multivariate auto-regressive (MVAR) model to the EEG data. The MVAR model is linear, so it will potentially miss out on the nonlinear relationships that are thought to be a part of EEG signals [155].

1.5.3 Mutual Information Analysis

In information theory, *entropy* is a fundamental quantity used to describe the uncertainty associated with a random system [254].

Definition 1.6. *For a discrete random variable X with possible states $\{x_i\}_{i=1}^N$ and associated probability mass function $P_X(x_i)$ the entropy is defined as*

$$H(X) = -\sum_{i=1}^{N} P_X(x_i) \log(P_X(x_i)).$$

Entropy can be used to describe the statistical dependence between two signals using *mutual information* (MI) [287]. If two signals have a statistical dependence, observing one signal reveals information about the other due to their common structure. The MI between two signals is the reduction in entropy caused by knowledge of the other. Mutual information analysis has been used to assess functional connectivity between signals in EEG recordings as early as 1985 [183, 300]. A benefit of using MI is that it can detect any type of statistical dependence, where as coherence is a measure based on linear dependence.

The ideas of MI analysis have been extended to measure directed connectivity using a quantity called *transfer entropy* [250], which measures the uncertainty associated with state space transitions. Analysis using transfer entropy has been successfully applied to electrophysiological data [213, 289]; however, state space reconstruction requires large amounts of data and is not always appropriate for EEG.

Information theoretic measures such mutual information and transfer entropy are able to detect non-linear dependency in EEG signals, however it can be difficult to get an accurate estimate of the probability distribution due to the non-stationarity of the signal.

1.5.4 Phase Synchrony

For as long as researchers have been doing quantitative spectral analysis, they have analyzed phase variables as a measure of time delay [2, 182, 201], however until the 1990s

amplitude measures such as coherence were the preferred measure of functional connectivity. Work done to understand the phase dynamics of coupled oscillators [223, 235] and new scientific models of how functional integration is achieved in the brain, using the theory of neural assemblies [43, 120, 119, 151, 178], have spurred increased interest in phase analysis for neural signals. Measures of connectivity using phase variables are based on the idea of *phase synchrony*, which is a consistent phase relationship between signals (see [286] for a review).

In order to identify periods of phase synchrony in EEG recordings, a time dependent *instantaneous phase* (see Section 1.1.1) is calculated using techniques based on the Hilbert transform [101, 113]. Due to the nature of EEG recordings, deterministic phase synchrony does not exist and an alternate definition of phase synchrony must be used. Several windowing methods have been proposed to measure phase synchrony; *phase locking values* (PLVs) are computed based on preferred values of the phase difference using entropy [270], the average phase difference over time [190] or trials [166, 232]. Windowing methods do not make full use of the time resolution of EEG signals; however, methods which use the phase difference derivative to identify phase synchrony [37, 93] are able capture the dynamic balance between synchronization and desynchronization thought to be an essential part of brain function [274].

There have been attempts to quantify the directionality of phase synchrony, models of weakly coupled oscillators have been used with time series data [52, 234, 259] and spike train data [159]. In [201] phase measurements based on the CSD are used to infer directionality in EEG, however there has not been work in identifying directed connections in the context of instantaneous phase shifting and locking events.

1.6 Modern Analysis - Source Reconstruction

Scalp topographies are able to localize the region of the scalp where an EEG signal is strongest, however the goal of EEG is to learn about brain function and the scalp is merely

a 2D projection of a 3D brain volume. Source reconstruction is the problem of identifying the location, orientation and magnitude of the primary electrical currents in the brain. With changing technology, this has become an active area of research.

To set up the source reconstruction problem, $Y_{s\times 1}$ is a vector of observed scalp potentials and $J_{3g\times 1}$ is the 3 dimensional primary current vector at g locations in the head, both measured in μV. A realistic range of values for s is 64-256 electrodes, while g should be in the 1000-2500 range to have high enough spatial resolution for measuring brain dynamics. This means $s << g$ and makes the source reconstruction problem underdetermined. Here I will review some of the literature on source reconstruction, and then these ideas will be revisited in Chapter 2.

1.6.1 The Forward Problem

The forward problem can be defined as determining the electric field potential that is generated by a given primary current density. Using Maxwell's equations (see Appendix A), the relationship between primary currents and scalp potential at a single time point is described by a Poisson equation [212]

$$\nabla^2 Y = \frac{\nabla J}{\sigma},$$

where σ is a conductivity parameter. For a discrete primary current density, the solution at a single point in time can be represented as a linear transformation with the *lead field matrix* (LFM) denoted K,

$$Y = KJ.$$

The LFM is a constant matrix which depends nonlinearly on the geometry of the problem as well as the conductive properties of the volume, a set of assumptions about these properties define a head model.

A simple head model is a uniform sphere with homogeneous and isotropic conductivity (σ is a constant). Based on experiments in [199] the assumptions of homogeneous and

isotropic conductivity are not realistic, however the single sphere model was originally used because there exists an analytic solution for LFM [41]. According to [7, 302] the single sphere head model is too simple and can give inaccurate results when calculating field potentials.

The Poisson equation can also be solved for inhomogeneous volumes [104] and has been used in multiple concentric spheres head model, where each sphere represents a part of the head (such as brain tissue, skull and scalp) with different conductivity [239]. A useful approximation was presented by Berg [23], showing that a model of M concentric spheres can be reasonably approximated by the sum of M appropriately scaled single sphere models.

Non-spherical head models also exist, imaging techniques such as structural MRI are used to create realistic boundary conditions which define homogeneous regions of the brain [255]. For these head models there is no analytic solution for the LFM, and numerical methods must be employed to calculate K [84].

In a spherical head, the potential can be described in terms of the contributions from the radial and tangential components of the primary current density [7]. One solution based on the structure of cortical columns, the primary currents are constrained to be perpendicular by setting the tangential component to zero [192].

1.6.2 The Inverse Problem

The forward problem is determining the electric potentials which are generated on the scalp due to specific generators within the brain. To perform source reconstruction, the corresponding inverse problem must be solved; that is, given observed scalp potentials, what are the underlying neural sources which generated those potentials.

Parametric Solution

Simple *dipoles* (point sources of primary currents) are able to provide a good description of EEG scalp data [36, 91]. The dipole method is a parametric solution to the EEG inverse problem which assumes that electrical activity in the brain is generated by a small number of point sources, each with six parameters describing the location and strength of the current in three dimensions.

The first implementations of the dipole solution were described in the early 1970's [139, 248], using just a single dipole at a specific point in time. In multiple dipole models, selecting the appropriate number of dipoles is an important aspect of an EEG source reconstruction. In is shown in [156] that a single dipole model correctly identifies the source of electrical activity due to blinking in participants with only one eye, however if both eyes are intact then the model incorrectly identifies the midpoint between the eyes as the generator. Additionally, if too many dipoles are fitted, they will tend to the same (correct) location, but with opposite current vectors which obscure the true magnitude of the source.

A number of extensions to the dipole method have been suggested (see [121] for a review), temporal evolution [246], subspace methods such as the *MUtiple SIgnal Classification* (MUSIC) and *Recursively Applied and Projected* (RAP-MUSIC) algorithms [193, 191], beamforming techniques [288] and computational artificial neural network approaches [1].

Nonparametric Solutions

Nonparametric solutions, also called distributed current (DC) solutions are another class of solution for the EEG inverse problem. The brain volume is discretized into many tiny elements called *voxels*. In [18], it is suggested that spatial and temporal resolution should be no worse than 5mm and 5ms respectively. In order to find a unique solution, regularization techniques are used to solve problems with additional constraints on the mathematical or physiological properties of the solution.

The earliest DC solution is the minimum norm estimate (MNE) [130, 131]. The use of a MNE corresponds to the minimum energy output required to generate the observed data, because of this it has a bias towards shallow sources and underestimates the depth of sources [216]. An extension of the MNE is *standardized LOw Resolution TomogrAphy* (sLORETA) [215], which uses a MNE normalized by an estimate of the source variance, allowing for better localization of deep brain sources. The estimate is based off a decomposition of sensor variance into independent contributions from source variance and additive noise. In noiseless simulations with a single dipole, sLORETA was shown to have zero localization error [215].

Another important type of DC solution is the weighted MNE. Different weights have been used for the EEG inverse problem, such as *LOw REsolution TomogrAphy* (LORETA, different from sLORETA) [216] and *Local AUto-Regressive Average* (LAURA) [70]. The LORETA solution seeks the 'smoothest' possible solution by minimizing the Laplacian weighted norm, and is able to correctly identify depth sources[216]. Although it provides more accurate source localization, the low resolution property of LORETA can sometimes blur together distinct but proximate sources.

There are other DC solutions that exist which use spatial regularization [33], spatio-temporal regularization [18], optimal resolution [69] and iterative multi-resolution solutions [114] (see [121, 214] for more detailed reviews).

Statistical Learning

A statistical approach to solving the source reconstruction problem is to use machine learning techniques [95]. The problems of find the function F which 'best' describes the relationship between two variables. The best function is the one which minimizes the expected loss for a given loss function. Here we are specifically interested in finding the function which best related the observed scalp recordings to the generating sources, subject to the constraint that the solutions satisfies the forward problems,

$$J = F(Y), \quad \textit{such that} \quad Y = KF(Y).$$

In situations where the number of potential predictors p is much larger than the number of observations N ($p >> N$), the standard approach is to include a regularization term to the loss function to encourage a sparse solution. Cost (loss, penalty) functions are used which encourage sparse solutions, such as the lasso (l_1 norm) and fused lasso [278, 279]. These regularization techniques usually involve a tuning parameter which must be set, this adds a model selection aspect to these methods. Cross-validation is a standard model selection technique which can be used to define the parameter [111].

Machine learning techniques are a growing field in neuroscience problems [172]. Applications of machine learning techniques to EEG source reconstruction, fused-lasso type regularization which results in sparsity and smooth current density [134, 133, 136]. An unsupervised learning algorithm which is applied to EEG is independent component analysis (ICA) [153]. There are many different applications of the ICA algorithm, popular choices are the FastICA [143] and Infomax [171] algorithms; there is also a popular Matlab toolbox called EEGLAB which includes a built-in ICA procedure [71].

Bayesian Solution

There are many possible solutions to the EEG inverse problem, each one representing different *a priori* assumptions about the mathematical or physiological properties of the current distribution. There have been studies which compare the accuracy of different solutions (for a review, see [121]), but there is no one solution which is uniformly better than all others.

A Bayesian statistical solution is a natural way to approach the EEG inverse problem, as many of the previously mentioned methods can be motivated using a Bayesian formulation [54, 53, 144]. The Bayesian approach also has a natural framework for model selection which allows for many different constraints to be compared, in a way that embodies the principals of Occam's Razor [283, 185].

A primary concern when using a Bayesian solution is that the posterior distribution can be intractable. Bayesian solutions have been used to incorporate complicated prior spatial information [220, 240] as well as spatio-temporal priors [18, 62], but these studies used *maximum a priori* (MAP) estimators due to intractable posterior distributions. Another option is to use Markov chain Monte Carlo (MCMC) sampling methods [103] to generate a random sample from the posterior distribution. These methods have been applied in Bayesian source reconstruction [13, 152, 247], however there is a high computational cost to generate large enough samples.

In [100], a Bayesian framework is developed for use in neural analysis, which uses hierarchical linear models to incorporate prior information and a *variational Bayes* (VB) approximation to the posterior distribution [12]. The framework has been applied to EEG inverse problems [158, 219, 243] and extended to include multiple prior distributions [97, 282].

1.7 Future EEG direction

The future of EEG connectivity analysis is in making inferences about functional integration at the cortex, as opposed to on the scalp. One option is to use source reconstruction techniques to find the cortical signals and then use traditional connectivity measures to identify regions of cortical connectivity. Source reconstruction techniques are able to localize EEG signals, however they can often distort the temporal properties of the signal and cause spurious patterns between signals [249].

Another option for measuring cortical connectivity is to use joint EEG-fMRI recordings in order to have combined high temporal and spatial resolution in recordings. The use of fMRI scanning introduces large electrical artefacts into the EEG signal, and modern signal processing techniques must be used in order to separate the brain activity from the noise due to the fMRI in the electrical recordings [285].

Another branch of EEG research which will benefit from changing technology is the development of a *brain computer interface* (BCI). Machine learning techniques are used to train BCI devices to respond to the electrical signals recorded from the scalp. Early BCI devices have been developed which assist in motor control, communication and gaming [188].

1.8 Summary

In this chapter, we have discussed some of the basic properties of EEG recordings. The takeaway is that EEG signals are the result of a highly complex biological process and have a low signal to noise ratio. Analysis of EEG recordings requires sophisticated methods to uncover the underlying brain signal and has previously been limited by the available technology.

The main problem discussed is estimating functional integration between neural regions from spatio-temporal patterns in EEG recording. Coherence has been the most popular measure in recent history, but it has limitations. New scientific models of how functional integration is achieved emphasize temporal coding in information transfer and encourage the use of phase variables in analysis. In Chapter 2 I introduce a novel measure of connectivity based on the existence of characteristic phase shift events which represent dissolution of neural assemblies. Shift events are modelled as a multivariate point process, and directed connectivity is assessed using the framework of Granger causality. One feature of the PSGC method which is specifically missing from the EEG phase shift literature is an objective way to identify phase shift events, in Chapter 3 I develop objective methods for phase shift identification, focusing specifically on the temporal resolution of each method.

A major limitation in estimating neural connectivity from EEG recordings is field spread due to VC, the instantaneous spread of electrical activity across the scalp. It is possible that a single source interacting with itself can generate spurious connectivity at distant locations, and even in the case of genuine connectivity, it is difficult to determine if the

true source of EEG activity is actually below the electrode. In Chapter 4, I investigate the effects of VC on the phase shift Granger causality (PSGC) analysis and the implications for inferring neural connectivity. Source reconstruction is one technique which may be use to overcome the limitations of VC, by identifying the primary currents within the brain which generated the observed scalp potential. This is a highly underdetermined problem, and different regularisations or constraints are used to find a unique solution. In Chapter 5, I explore the ability of several source reconstruction techniques to accurately localize the source of a phase shift event.

Chapter 2

Phase Shift Granger Causality

Electroencephalogram (EEG) signals across two sites are said to be functionally connected with one another when their interaction has a physiological relevance [35], such as during periods of synchronized phase relations [119, 286]. However, desynchronization of neural assemblies is a necessary aspect of information processing [277]. Traditional measures of EEG functional connectivity do not consider the relationship between synchrony and desynchrony, focusing rather on an overall measure of synchronization strength [205, 201]. In this chapter, I introduce a novel method of measuring directed connectivity from spatio-temporal patterns of these phase shift events by applying the framework of Granger causality (see §1.5.2 or Appendix E).

The current Chapter is a preliminary demonstration of this new measure, applied to EEG recordings from 18 adolescents during a rest, and both auditory and visual vigilance tasks. The properties of this method are further studied in Chapter 3 and Chapter 4, before a final validation is given at the end of Chapter 4. The new measure, *Phase Shift Granger Causality* (PSGC), is able to clearly distinguish between the resting task and the active tasks. The latter have higher rates of connectivity overall, and specifically more long range connections. The resting task appears to activate more localized neural circuitry while the active tasks appear to increase communication across several neural regions involved in vigilance tasks, this is consistent with the idea that the increased complexity of the

vigilance task requires additional resources to complete. The vigilance tasks also showed significantly higher clustering coefficients than the resting task, one property associated with small world network dynamics [295].

2.1 Introduction

The coordination of multiple functionally segregated brain regions to achieve a unified response is defined as *functional integration*. For many types of cognitive processing the integration of multiple regions is required [38, 45, 281]. For example, consider the act of writing; this requires the visual system to see the writing surface, memory and language to know what to write, and the motor system to tell the hand how to write. It is hypothesized that basic tasks can be mediated by the formation of neural assemblies [35], which are transient networks of functionally connected neurons.

Bidirectional connections in neural assemblies facilitate stronger interactions among neurons within an assembly via a feedback mechanism [88, 119, 261, 280]. Due to the common orientation of neurons within a cortical column, and the relative closeness of the cortical surface to the scalp, such activation patterns are ideally suited to be measured from EEG scalp recordings (see §1.3). Transfer of information within neural assemblies that is accomplished by these amplified cycles of activation will manifest in EEG recording as an oscillations within a specific frequency. This is idea is reinforced by the observation of distinct oscillatory components in EEG recordings [25, 65, 237, 47].

Graphs are mathematical constructs which can be used to study connectivity, with neural regions represented as nodes, and pairs of nodes have an edge between them if they are functionally connected. Complex systems often have graphs with structured connectivity patterns, such as the small-world [295] and scale-free [21] graphs. Both types of networks have been observed in the brain; structural networks show small-world behaviour [263], while transient functional networks display both small-world and scale-free behaviour [260]. Anomalous functional brain networks have been associated with patho-

logical conditions such as schizophrenia [176] and epilepsy [175]. The estimated networks depend the specific measure of functional connectivity which is used, the appropriate measure depends on the time-scale of the underlying process and the modality of brain imaging.

The literature further distinguishes between *functional connectivity*, a symmetric measure of statistical dependence between sources of neural activity, and *effective connectivity* which describes the influence that one neural system has over another [99] which is potentially asymmetric. Coherence has been used to measure functional connectivity at specific frequencies [164, 205, 201] and estimate network structure [194, 271] in EEG recordings, but this method is limited. Measures of coherence are based on an average of time segments which are one to two seconds in length and this level of temporal resolution might not fully capture the moment to moment changes in connectivity that are hypothesized to exist [262, 274, 277].

One limitation of coherence, the inability to measure direct connectivity, may be overcome using statistical techniques based on the concept of *Granger causality* [116] (see §1.5.2 or Appendix E). This idea of causality centres around predictability, and more specifically whether knowledge of one signal can improve the prediction of another. Measures of Granger causality are asymmetric and can be used to infer directed connectivity between neural signals. Granger's original paper introduced a statistical test for causality between two processes based on a bivariate auto-regressive (AR) model [116]. The test has a convenient spectral representation called directed coherence (DC), which decomposes the traditional coherence functions into two asymmetric components of directed connectivity. Originally used to study economic time series, the method was applied in neuroscience over a decade later [241]. An extension of Granger's method is called *partial directed coherence* (PDC), which provides a test for causality that takes into account the effect of other signals [15]; the method has been applied to many neuroscience problems in recent years [11, 245, 268]. Granger's original test for causality assumes a stationary AR process, although there is evidence that this condition may not hold in EEG recordings longer than 12 seconds [55]. This has led to the development of time dependent measures of Granger causality [9, 140, 189].

2.1.1 Phase

In order to account for temporal coding in frequency analysis, the phase of a signal can be analyzed. While there are a number of ways to define phase when dealing with noisy or chaotic signals [31, 223], *instantaneous phase* is a term used to describe the time-dependent phase of such signals [48, 227]. In this work, instantaneous phase is defined in the *Hilbert transform* sense, a concept first introduced by Gabor in 1946 [101] (see Section 1.1.1).

One important method of describing functional connectivity based on phase variables is through phase synchrony [286]. Two oscillators are said to display phase synchrony if they have a constant difference in phase over time. The presence of noise in a signal can lead to *phase slips*, small deviations from perfect synchrony [31, 270]. Thus in order to identify phase synchrony in the presence of phase slips (i.e. noise), an alternative definition is required. One option is an entropy-based method for quantifying the average strength of phase synchrony over time in noisy systems [270]. An alternate method of identifying phase synchrony proposed by Breakspear [37] uses the *phase difference derivative* (PDD) to identify periods of *phase locking* with high temporal resolution.

An interesting property of brain dynamics is the ability for spatially distant neurons to synchronize and form neural assemblies over both short and long distances [274, 272], leading to a two-compartment model for neural interaction [276]. Such synchronization could be physiologically facilitated by the thalomocortical loop - a subcortical tract of neural fibers which connects the thalamus to several regions of the cerebral cortex [151, 178, 265]. In models of coupled oscillators, synchronization can be accomplished with a brief low energy pulse (desynchronization or *phase shift*), which resets and then synchronizes the oscillators [221]. The occurrence of a phase shift followed by a period of phase locking has been called a *phase reset* (PR), and this dynamic balance between synchronization and desynchronization is thought to be an essential aspect of normal brain function [277].

Researchers have attempted to quantify the direction of information flow in phase synchrony. Models of weakly coupled oscillators have been used with time series data [52, 234, 233, 259] and with spike train data [159]; however, these methods are not optimized for use on EEG recordings. Phase slope index (PSI) is a method of inferring the direction of phase relationships [202], which has been successfully applied to EEG recordings. In contrast to this work, estimates of PSI are based on the mean phase (in the Fourier sense) over 1-2 seconds of data and quantify the consistency of the phase difference over that period.

In this chapter, a measure is outlined for estimating directed network connectivity in EEG recordings, based on patterns of phase shifts events. To establish the direction of functional connections, a measure is developed based on the ideas of Granger causality. Specifically, a *Phase Shift Granger Causality* (PSGC) measure is introduced based on the theory of point processes [59], an approach which has recently been successfully applied in the study of spike trains [160].

2.2 Methods

A sample EEG recording from regions F3, C3, P3 and O1 can be seen in Figure 2.1. These data will be used to illustrate the methods of analysis used in this paper. The estimated instantaneous phase in the 8-10 Hz band of the signals from the sample data are shown in Figure 2.2. The plot shows periods where the phase is approximately constant as well as sudden jumps in the instantaneous phase.

2.2.1 Phase Synchrony

In the study of periodic oscillators, two oscillators with the same frequency are said to display phase synchrony if there is a constant relationship between phases [31]

$$\Phi_t = \left| \phi_t^{(x)} - \phi_t^{(y)} \right| = k, \quad \forall t.$$

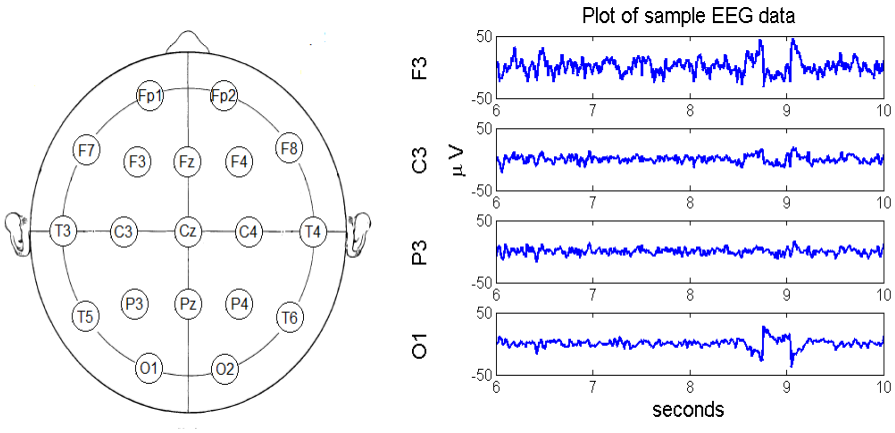

Figure 2.1: Left: A visualization of the international 10-20 system for electrode placement, originally published in [147]. The present study only uses regions selected from the left hemisphere (Fp1, F7, F3, C3, P3, T3, T5, O1) and the right hemisphere (Fp2, F8, F4, C4, P4, T4, T6, O2), and excludes all midline sites. Right: A sample of 4 seconds of resting EEG, recorded from channels F3, C3, P3 and O1.

To identify periods of synchrony in the presence of noise, such as in PR analysis, Breakspear's method [37] defines two signals as phase locked when their phase difference derivative is below a threshold value c

$$\left| \frac{d\Phi_t}{dt} \right| < c, \quad \forall t.$$

Phase locking has been used as a synonym for phase synchrony; here it is used to describe this specific measure of phase synchrony.

A prolonged period of phase locking followed by a brief desynchronization, or *phase shift* is described as a phase reset. The phase shift is hypothesized to represent the dissolution of a neural assembly and the accumulation of resources for the next task. Phase reset analysis involves the average locking and shifting durations; these variables have been shown to be correlated with brain development [277]. In a collaboration with the Cognitive and Affective EEG lab at Brock University, we have been able to replicate these results in the

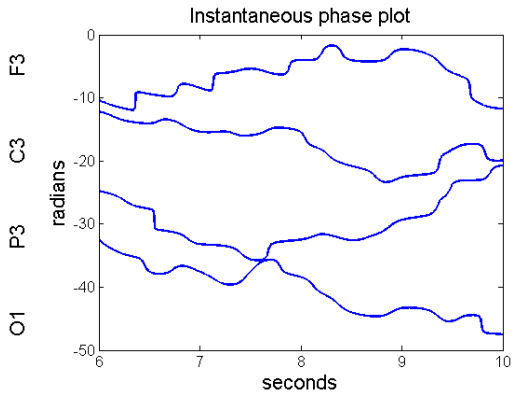

Figure 2.2: Instantaneous phase estimates of the 4 channels from Figure 2.1. There are two important features in these plots, a slight negative trend and sudden jumps in the phase. The slight negative trend of the phase suggests that there is power in the signal due to unrelated sources or noise, which occurs below the centre of the frequency band. The sudden jumps are phase shift events. Many of the shift events are obvious, but there are some places where it is difficult to tell if there is a phase slip or just noise in the signal (for example, at the O1 signal just before 9 seconds). This ambiguity is dealt with using a thresholding technique (see Section 2.2.2).

context of anxiety and aggression in adolescents [167].

2.2.2 Phase Shifts

When the estimated phase difference derivative first exceeds the threshold value c in phase reset analysis, there is said to be a phase shift [277]; this process necessarily involves a shift in the phase of one or both of the corresponding oscillators. Here we propose to analyze these characteristic phase shift events outside the context of phase synchrony across channels. Rather than identifying shifts in the phase difference between two channels, phase shifts can be identified for each individual channel. If ϕ_t represents the instantaneous phase

of a time series, there is said to be a phase shift at time t if

$$\left|\frac{d\phi_{t-1}}{dt}\right| < c, \quad and \quad \left|\frac{d\phi_t}{dt}\right| > c. \tag{2.1}$$

In previous studies, a simple visual inspection has been used to determine an appropriate value of c [37, 277], this includes an unwanted subjective aspect to the analysis. Note also that the appropriate range of values for c depend on properties of the signal such as the sampling rate, the filters applied and the frequency band of interest and therefore may differ from study to study. Initially, for illustration, we select the threshold value in a subjective manner; however, the development of an objective alternative is the subject of Chapter 3.

In this context, patterns in phase shift events can be analyzed in a multivariate model of EEG signals, taking into account events from across the scalp. Figure 2.3 shows the instantaneous phase derivative of the example data, a phase shift event is defined to occur at the local maximum of the phase difference derivative.

2.2.3 Phase Shift Granger Causality

A point process is a model which describes the occurrence of random events localized in either time, space or both. There are a number of equivalent ways to describe a point process; for a review of these and some other properties see [59]. This document will only be concerned with temporal point processes, corresponding to events on the positive real line.

A continuous time point process is a random measure on a completely separable metric space which is restricted to $\mathbb{Z}^+$ [59]. More informally, a point process can be described by an integer-valued random step function N(t), defined to be the number of events which occur in the interval (0,t]. A set of observed data from a point process over a fixed interval [0,T], called a *sample path*, is an increasing sequence of numbers $0 < t_1 < t_2 < ... < t_K < T$,

which defines the time of each event, and where K is a discrete random variable with support over the non-negative integers.

In the current context, a sample path for the point process $N_i(t)$ (i=1..S) describes the time of each phase shift event identified from the i^{th} signal of the EEG recording. A sample point process representation of phase shift events is shown in Figure 2.4. If the number of events in successive intervals are independent, the probability structure of a point process model can be completely described in terms of an inhomogeneous Poisson process with intensity function

$$\lambda(t) = \lim_{h \to 0} \frac{P(N(t+h) - N(t) = 1)}{h}.$$

When the number of events in successive intervals are not independent or depend on other variables, a *conditional intensity function* (CIF) can be used to describe the distribution of the process, which is a stochastic process of the form,

$$\tilde{\lambda}(t) = \lim_{h \to 0} \frac{P(N(t+h) - N(t) = 1 \mid \mathbb{H}_t)}{h}.$$

Here the term $\mathbb{H}_t$ is the complete history of the process up to but not including time t. The history can also include information from other processes or various exogenous variables. The CIF is interpreted as

$$\tilde{\lambda}(t)dt \approx E(N(t+dt) - N(t) \mid \mathbb{H}_t).$$

Exactly what information is in $\mathbb{H}_t$ and how it is incorporated into the CIF depends on the context of the problem.

For a recording with S channels, let $N_i(t)$ (i=1..S) define the phase shift point process of the i^{th} signal. Following Kim [160], the history $\mathbb{H}_t$ will be the number of events R(t) in M consecutive non-overlapping windows of length W that occur just prior to t. The history will also contain similar information for each signal in the recording,

$$\mathbb{H}_t = \{R_{i,m} | i = 1..S, \quad m = 1..M\},$$

where

$$R_{i,m}(t) = N_i(t - (m-1)W) - N_i(t - mW).$$

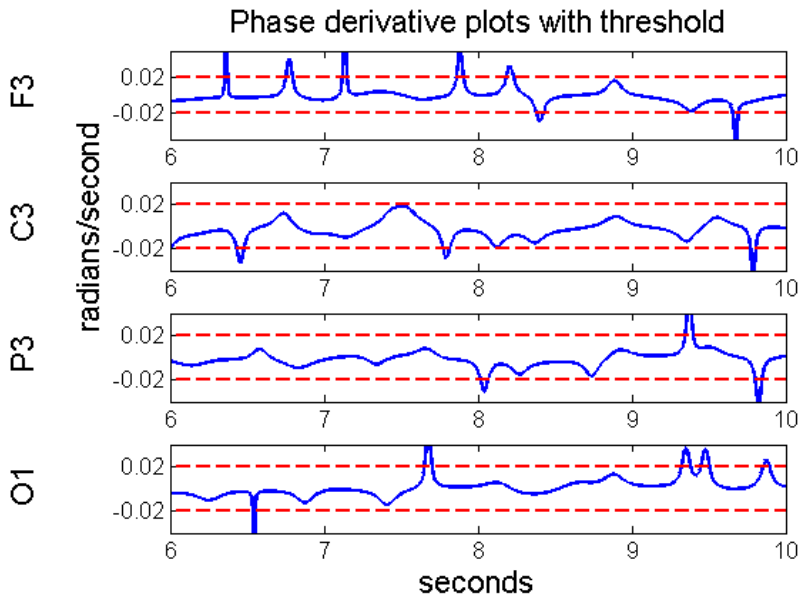

Figure 2.3: Plots of the phase derivative of the four instantaneous phase signals in Figure 2.2, including a dashed line indicating a threshold value of $c = 0.02$. Notice that the deviations of the phase derivative of channel O1 just before 9 seconds were not large enough to be classified as a phase shift.

The values $R_{i,m}(t)$ are then be used to define a log-linear parametrization of the CIF for process N_j,

$$\log \lambda_j(t) = \gamma_{j,0} + \sum_{i,m}^{S,M} \gamma_{i,m} R_{i,m}(t).$$

The parameters $\gamma_{i,\cdot}$ represent the effect that N_i has on the intensity of N_j.

Given a sample path $\{t_1, t_2, ..., t_{N_j(T)}\}$ for channel N_j, the likelihood function for this model is given by

$$L_j(\vec{\gamma}) = e^{-\int_0^T \lambda_j(u)du} \prod_{i=1}^{N_j(T)} \lambda_j(t_i).$$

If the time interval is discretized into small bins of length Δ, and Δ is chosen such that

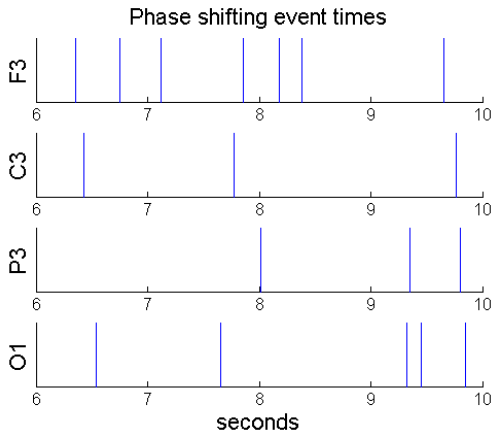

Figure 2.4: A visualization of the point process of phase shift events identified in Figure 2.3. Spatial and temporal patterns in these events are further analyzed using Granger causality in Section 2.2.3

the probability of two or more events in a single bin is negligible,

$$P(N_j(t+\Delta) - N_j(t) \geq 2) \approx 0,$$

then based on the Poisson assumption, the likelihood function can be approximated as a product of the probability functions for independent non-identical Bernoulli trials [59]. We define $dN(t_k)$ as the number of events in the k^{th} bin, that is

$$dN_j(t_k) = N_j(t_k) - N_j(t_k - \Delta), \quad k = 1..K,$$

then,

$$L_j(\vec{\gamma}) = \prod_{k=1}^{K} (\lambda_j(t_k)\Delta)^{dN_j(t_k)} (1 - \lambda_j(t_k)\Delta)^{1-dN_j(t_k)}. \tag{2.2}$$

This approximation is particularly useful because it leads to a convex likelihood function, allowing easy access to the maximum likelihood estimates of the parameters. A standard Newton's method was used to find the maximum likelihood estimates for the parameters; for computational reasons a lower bound on parameters of -10 was included in the estimation, otherwise the numerical optimization algorithm may diverge to negative infinity.

Parameter steps are reduced by a factor of 5 to prevent overshooting critical points, the update equation for the algorithm is given by

$$\lambda^{(i+1)} = \lambda^{(i)} + \frac{1}{5} I(\lambda^{(i)})^{-1} S(\lambda^{(i)}),$$

where I and S are the information and score functions corresponding to the approximate likelihood function.

An important aspect of Granger causality measures is a test of the hypothesis H_0: 'N_i does not Granger cause N_j'. By comparing the full likelihood for N_j with a reduced likelihood with the history of N_i removed ($R_{i,\cdot}(t)$), we can test whether the information in N_i causes a reduction in the variance of the prediction error. To determine the effect of the signal N_i on the optimal predictor for N_j, a reduced likelihood function is defined excluding the parameters $\gamma_{i,\cdot}$ from the CIF

$$\log \lambda^*_{i,j}(t) = \gamma_0 + \sum_m \sum_{q \neq i} \gamma_{q,m} R_{q,m}(t),$$

$$L^*_{i,j}(\vec{\gamma}) = \prod_{k=1}^{K} (\lambda^*_{i,j}(t_k)\Delta)^{dN_j(t_k)} (1 - \lambda^*_{i,j}(t_k)\Delta)^{1-dN_j(t_k)}.$$

The likelihood ratio of the full and reduced models is used to defined PSGC $\Phi_{i,j}$, an asymmetric measure of connectivity from N_i to N_J,

$$\Phi_{i,j} = -2 \log \frac{L^*_{i,j}}{L_j}. \tag{2.3}$$

To assess the significance of the PSGC measure, a natural hypothesis test exists in the form of a likelihood ratio test (LRT). Under the null hypothesis, for large values of T, Φ follow an approximate chi-squared distribution with M degrees of freedom,

$$\Phi_{i,j} \sim \chi^2_M. \tag{2.4}$$

In a simulation study to be described in Chapter 4, a time of $T = 5$ minutes was sufficient for accurate results from the LRT.

This provides a test for Granger causality from N_i to N_j within the framework of a multivariate model of the signals, meaning that the test is able to account for common information contained in multiple signals. If N_i is estimated to Granger-cause N_j, then knowledge of past shift events in N_i improves prediction of future shift events in N_j. In other words, if there is a shift at N_i, then it will either increase or decrease the probability that N_j will shift in the imminent future. It can occur that a Granger causal relationship is estimated in both directions; in such situations it is not possible to determine if one of the signals is actually 'leading' the other, either due to insufficient temporal resolution, or bidirectional coupling.

As an additional method of statistical validation for the PSGC measure, we test the hypothesis that the mean net Granger causality is zero, using the method of time-reversed surrogate data [135]. By time-reversing the data ($x_i^* = x_{N-i}$) a surrogate dataset is created which maintains most properties of the signal, while reversing the temporal structure and hence the Granger causality. A statistic is calculated to by looking at the difference between net causality scores in the forward (Φ) and time-reversed (Φ^*) data,

$$d = (\Phi_{i,j} - \Phi_{j,i}) - (\Phi^*_{i,j} - \Phi^*_{j,i}). \tag{2.5}$$

If there is asymmetric coupling, then terms $(\Phi_{i,j} - \Phi_{j,i})$ and $(\Phi^*_{i,j} - \Phi^*_{j,i})$ will have opposite signs, resulting in a large value of d, while for a bi-directional coupling both terms should be equal, resulting in a value of d close to zero. The null hypothesis of zero net connectivity ($\mu_d = 0$) can be tested using a standard t-test or non-parametric alternative. The test for net GC is not sensitive to so-called weak asymmetries [135] such as those introduced by volume conduction (VC). Since the degrees of freedom (M) may change for each dataset (participant), PSGC scores we standardized before applying the analysis by converting to p-values, and then an inverse transformation was used to generate values from a χ^2_4 distribution. We will subsequently refer to this second statistical test as a net Granger Causality (netGC) analysis.

2.3 Functional Application

Using the PSGC method, we can estimate directional phase relationships in the 8-10 Hz band of the electrocortical signal relations across scalp sites. The data in this study derive from 18 recordings and three conditions: a resting condition, an auditory vigilance task and a visual vigilance task (see §1.1.3).

Predictions: The relatively complex attention demands of the vigilance tasks require functional connection of networks associated with both the frontal and temporal/parietal attention systems [24, 42, 238, 256]. In contrast, the resting condition should engage in more localized connections given that there are no complex demands on the task. We predict that these differences will results in higher levels of network connectivity, as measured by the total number of connections, as well as increased average clustering coefficients. Additionally, we separate the edges into short and long connections (adjacent scalp sites versus those separated by more than one location) and expect to see an increase in long range connections for the vigilance tasks.

2.3.1 Application Details

Based on the results of previous analysis of the EEG tasks using phase reset measures and their association with individual difference variables [167], the frequency band of interest for these analyses was 8-10 Hz. The instantaneous phase of each channel was calculated by shifting the centre of the band of interest to 0 Hz and then putting the result through a 1 Hz low pass Butterworth filter (implemented using Matlab *butterworth()* and *filtfilt()* functions).

For the point process model, the length of the history windows W must be chosen so that the time scale is appropriate in the context of behavioural analysis. For this study, a value of $W = 100$ ms was used. This value of W is approximately the length of a complete cycle at the frequency band of interest. For each recording (real and simulated data) and

each channel, model selection was performed to select the optimal order M for the log-linear parametrization. The optimal value was determined by calculating the Akaike information criterion [4] for a range of values of M (1 to 6) and selecting the one which minimized the criteria. This process creates a balance between goodness of fit and overparametrization of the model. The optimal values of the order parameter M ranged from 3 to 5, corresponding to a total history length of 300-500 ms which is enough time to capture a decision-making process.

Results for both statistical analyses were corrected for multiple comparisons using the Holm-Bonferroni procedure [141] to give a familywise error rate of $\alpha = 0.05$ per hemisphere. For individual networks estimated using the LRT, corrections are applied based on (8 x 7) = 56 comparisons per hemisphere and (8 x 7)/2 = 28 multiple comparisons for the netGC analysis.

2.3.2 Results

In order to simplify the demonstration, the analysis was restricted to within-hemisphere connections. For each participant and each ordered pair (i,j) ($i \neq j$), the PSGC $\Phi_{i,j}$ was calculated and the LRT is used to test the hypothesis that there is no Granger causality from N_i to N_j. Figure 2.5 shows the connectivity patterns for each of the six task (3) by hemisphere (2) combinations. In these figures, each cell of the matrix represents the Granger causality from one neural region to another. The i,j element of the matrix represents PSGC from N_i to N_j. The shade of the cells represent the number of participants who had significant connectivity in the specified pathway. The i,j element is white if causality exists in every participant, black if there is causality in zero participants and shades of grey represent intermediate results. The diagonal entries are crossed out to emphasize that there is no analysis on these locations. From the matrices, it can be seen that there were higher levels of connectivity during the active tasks (auditory, visual) than during the resting condition. Different groups of connections were active during the resting task compared to the two active tasks, for example between regions F8, T4 and T6 (right hemisphere), and regions F7, T3 and T5 (left hemisphere). There is no attempt made to

interpret similarities or differences between hemispheres.

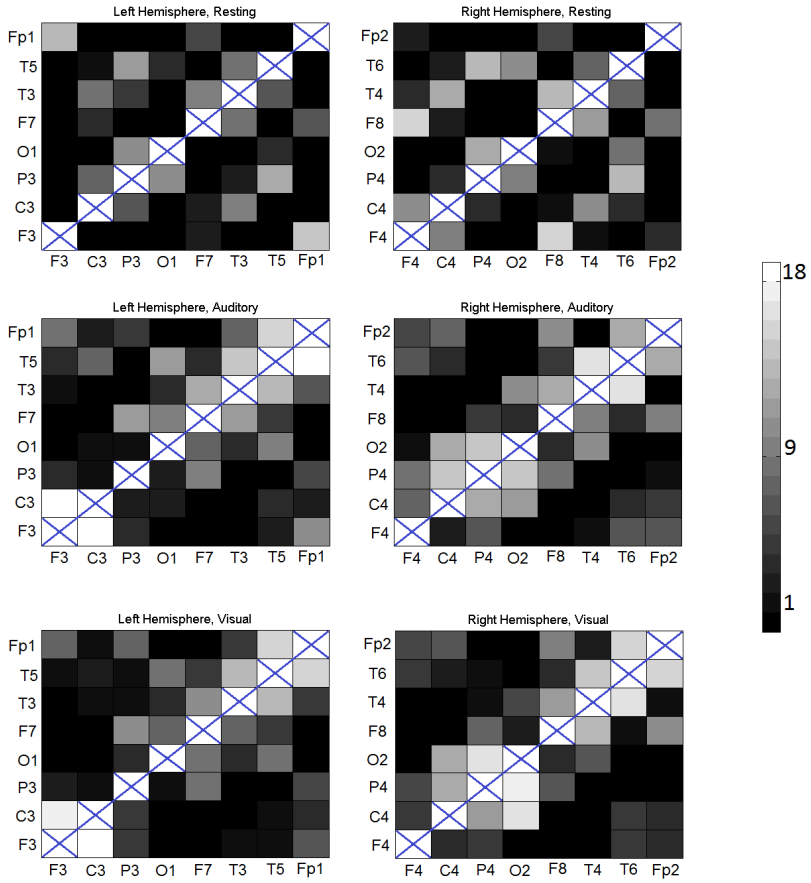

Figure 2.5: Results of the PSGC analysis between electrode sites in 18 EEG recordings. The i,j cell of the matrix represents the PSGC from region i to region j. The grey scale (see colourbar) represents the number of participants where there is significant connectivity. A black cell means no participants had significant connectivity, while a white square signifies information transmission in all 18 participants. Results among left hemisphere locations are on the left, right hemisphere locations on the right. The three rows correspond to resting (top), auditory (middle), and visual (bottom).

Sample connectivity maps from two participants can be seen in Figure 2.6 (connectivity maps for all participants can be found in Appendix F). In the resting condition, the LRT finds many short range (neighbouring site) connections, while both the auditory and visual tasks show more long range connections. The connections in the vigilance tasks are predominantly anterior-posterior, connecting the prefrontal sites to the parietal sites, instead of lateral-medial. Based on previous literature, it was our hypothesis that there would be behaviourally relevant differences in this 8-10 Hz frequency band, although there were no hypotheses on the nature of these differences. A post-hoc explanation for these differences is that they may reflect engagement of the frontal and posterior attention systems in the two vigilance tasks [293]. The connectivity matrices do not show any consistent asymmetries (consistent with the netGC analysis below); however, looking at individual networks suggests that asymmetries may exist but are not consistent across participants due to individual differences.

2.3.3 Mean Effects

In addition estimating network connectivity for each recordings, we also analyzed network patterns for differences between tasks which are consistent across many participants. To identify asymmetric connections which were active across many participants, the netGC analysis was used; only two significant asymmetries were identified, a short range prefrontal scalp connection in the resting task ($F7 \rightarrow F3$) and a connection leaving the scalp regions above the visual cortex ($O2 \rightarrow C4$) in the visual condition. The lack of asymmetric connectivity at the group level, when compared to the individual networks, suggests that either the results for individual networks are largely the result of a bidirectional coupling, or that there are individual differences in the size and direction of asymmetric connectivity.

Individual networks were additionally quantified using the network average of local clustering coefficients, a graph theoretical measure which has been used in describing 'small-world' networks in functional connectivity [263, 295]. The local clustering coefficient measures the degree to which neighbors of a node are interconnected. The average clustering

coefficients were used as the dependent variable in a 2 (hemispheres) x 3 (conditions) repeated measures ANOVA, with p-value corrections for sphericity using Greenhouse-Geisser method [122]. The resting task ($\hat{\mu} = 0.095$, $\hat{\sigma} = 0.027$) had significantly lower clustering coefficients than the auditory ($\hat{\mu} = 0.216$, $\hat{\sigma} = 0.034$) and visual tasks ($\hat{\mu} = 0.198$, $\hat{\sigma} = 0.038$), $F(2, 34) = 4.858$, $p = 0.017$.

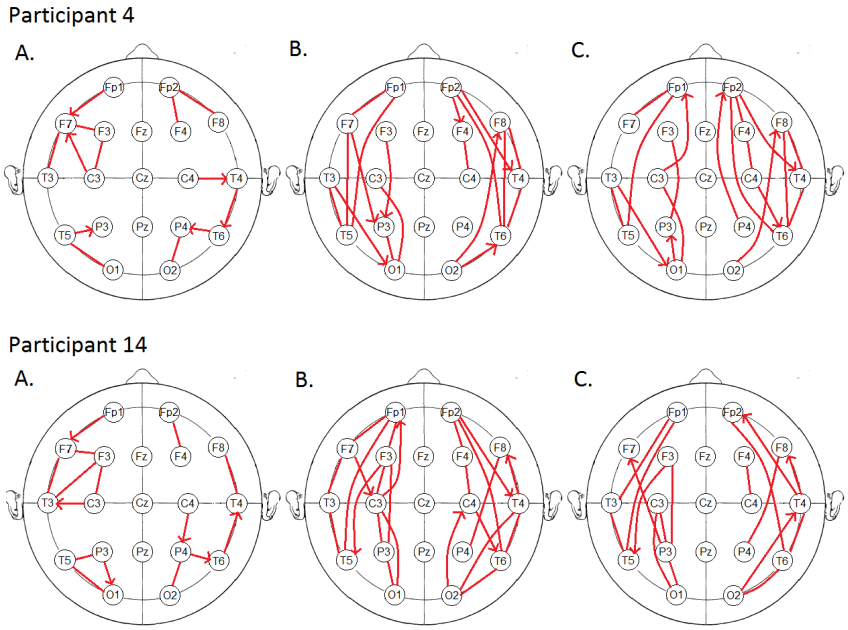

Figure 2.6: Maps of the estimated network connectivity during the A: resting, B: auditory, and C: visual tasks, for the χ^2 analysis of two individual participants. In the resting task we see many short range connections but relatively few long range connections. In contrast, the active tasks show many connections over longer distances, primarily anterior-posterior, connecting prefrontal sites to parietal sites. Asymmetries are observed in these individual networks; however, the direction is not consistent between individuals. This pattern was consistent across all participants and is further quantified in Figure 2.7.

We performed a further test of the PSGC effects by examining the directed connection

patterns as a function of the three conditions across participants. To follow up on the descriptive findings outlined above, we focused on short versus long connections. The number of significant directed connections per person was the dependent variable in a 2 (hemispheres) x 3 (conditions) x 2 (connection lengths) within-subject analysis of variance, with p-value corrections for sphericity [122]. The right hemisphere sites ($\hat{\mu} = 7.1$, $\hat{\sigma} = .358$) demonstrated significantly more overall connections than the left ($\hat{\mu} = 6.0$, $\hat{\sigma} = .349$), $F(1, 17) = 7.95$, $p = .012$. Similarly, the vigilance tasks elicited more connections ($\hat{\mu} = 7.6$, $\hat{\sigma} = .550$ and $\hat{\mu} = 7.0$, $\hat{\sigma} = .433$ for the auditory and visual tasks, respectively) than did the resting condition ($\hat{\mu} = 5.2$, $\hat{\sigma} = .373$), $F(2, 34) = 9.03$, $p = .001$. There were significantly more short ($\hat{\mu} = 8.1$, $\hat{\sigma} = .368$) than long ($\hat{\mu} = 5.1$, $\hat{\sigma} = .313$) connections, $F(1, 17) = 89.6$, $p < .001$. Of particular interest was the Task x Connection Length interaction, which was highly reliable, $F(2, 34) = 145.5$, $p < .0001$, with the mean values illustrated in Figure 2.7. To examine the reliability of this pattern further, we constructed for each participant a metric reflecting the extent to which the resting condition was associated with more short connections than the active tasks and the active tasks more long connections than the resting condition. This yielded totals for the left hemisphere and right hemisphere for each participant. Each hemisphere of every participant followed this general classification (range = 1 to 20, with any value above zero indicating this general pattern), indicating extremely high reliability of the pattern illustrated in Figure 2.7.

2.4 Discussion

This chapter introduces a novel method for analyzing neural connectivity from EEG recordings. The PSGC measure is a phase-based method which identifies spontaneous desynchronizations called phase shifts, and assuming such events are identified, they are modeled as a multivariate point process. By applying the ideas of Granger causality, an asymmetric measure is produced which takes into account the effect of all other signals and can be used to infer effective connectivity between neural regions. As a phase-based measure, PSGC is able to capture nonlinear relationships which are not detectable by traditional coherence analysis. By using instantaneous phase to identify phase shift events, it takes advantage of the full temporal resolution of EEG recordings because there is no need to average over

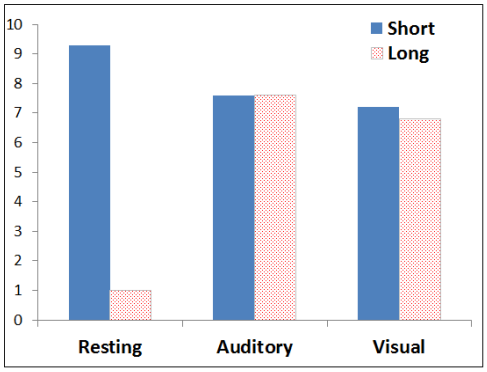

Figure 2.7: The significant Task x Connection Length interaction, demonstrating an increase in longer connections and a reduction in shorter connections during the auditory and visual vigilance tasks compared to the resting condition.

segments.

I have presented two different statistical techniques for analyzing the results of PSGC, a LRT for Granger causality in individual recordings (see Equation 2.3) and a test for zero mean net Granger causality (see Equation 2.5) which uses time-reversed surrogate data [135]. The netGC analysis is the preferred test for identifying asymmetric connectivity at the group level; it is designed to maximize the statistical power to identify asymmetries. The LRT identifies both asymmetric and bidirectional connections, allowing the compete networks to be estimated and then summarized by graph theory metrics. A benefit of the LRT is that it provides statistical validation at the individual level, whereas the netGC analysis is restricted to group results which assume commonality of a specific directed connection across individuals. Given the multiple processes involved in all experimental tasks and the mixture of signals at the scalp, we do not expect group commonality as reliably as we see individual effects.

The primary result of interest supports the prediction concerning connectivity during complex attention versus a resting state: the number of significant short connections

during the resting state exceeds those during the vigilance tasks, but of more importance, the longer distance connections dramatically increase during the vigilance tasks. This is consistent with the expected connectivity given the neurophysiology of attention [24, 42, 238, 256]. There were in addition some main effects that also confirmed expectations. The active tasks had significantly higher clustering coefficients than the resting task, a feature associated with small-world networks [260]. The right hemisphere showed more significant instances of connectivity than the left, consistent with findings of more white matter in the right than in the left hemisphere [124]. Similarly, there was an increase (albeit minor) in overall connectivity during the active tasks. The finding of more significant instances of connectivity for the short (across adjacent sites) compared to the long (across more distant sites) may be due to passive transmission being greater over short distances, and is only of interest when interacting with task requirements, as was found in our study, with this pattern present for every participant. Converging evidence from these measures of network connectivity strongly suggest differing patterns of neural connectivity between resting and active tasks.

2.4.1 Limitations and Future Work

Based on reviewer feedback received on the PSGC method, there were several areas which have been targeted to be addressed in the remaining chapters.

A limitation of PSGC is the subjective choice of the threshold value c. Sensitivity analysis of the time between phase shift events (phase locking intervals) has shown that both the average interval length [167] and the shape of the interval distribution [37] are not sensitive to the specific value of c. In Chapter 3 I investigate objective methods for estimating the value of c such that it effectively identifies the phase shift events.

The primary result of interest involves the differences in the estimated connectivity networks between tasks. A concern is that VC (see 1.6.1) may confound the estimation of connectivity from scalp recordings [249], and whether these differences are actually a

result of changes in neural connectivity or if it is some other feature of the recordings that has changed. In Chapter 4 I present a simulation study which elucidates the effects of VC on the PSGC method, and allows for more informed conclusions about the results of the PSGC analysis.

A natural extension of the PSGC method is to evolve it to work on the results of source reconstruction techniques, to further reduce any effects of VC and increase spatial resolution for localizing of neural regions. Proceeding in this manner should be done with caution, as at this time it is not clear whether such inverse solutions are able to properly maintain phase relations. In Chapter 5 I evaluate several methods of source reconstruction on their ability to accurately localize phase shift events.

Another potential area of improvement is the point process model for phase shift events. This was originally planned to be one aspect of the thesis, but the other topics were investigated instead based on reviewer feedback. Some ideas for this include a refractory period after shift events, or a different distribution which is generates the heavy-tailed inter-shift-interval distributions observed in the data (see Chapter 3). This could also include updates to what is included in the history $\mathbb{H}_t$, such as stimuli onset or response time events.

Chapter 3

Detection of Phase Shift Events

This chapter develops statistical methodology which is suitable for tackling the problem of (instantaneous) phase shift identification from observable time series, such as in the phase reset (PR) and phase shift Granger causality (PSGC) analyses discussed in the previous chapter. Such change points, or *phase shifts*, can be markers of information transfer in complex systems; their analysis occurring in geology, biology and physics, but most notably in neuroscience. Since these phase variables are computed from complex systems which often include significant amounts of noise, we focus on developing methods which are robust to such noise, but also have the temporal resolution necessary for many application areas. Two non-parametric approaches are developed for this problem: the cumulative summation (CUSUM) and phase derivative (PD) estimators. A system of weakly coupled Rössler attractors provides an application in which there are high levels of systematic and time-dependent noise. Shift identification is also performed on beta-band activity from electroencephalogram (EEG) recordings of a visual attention task, an unsupervised application which requires high temporal resolution.

3.1 Introduction

The analysis of phase change points has applications in many fields. In neuroscience, phase shift behaviour is present on small-scale (cellular) analysis of spike bursting data [146]. It has also been applied to other complex biological oscillators, such as the Circadian rhythm [187] and in cardiac/respiratory systems [20]. In geology, directional (circular) data is used to identify faults, pinchouts and other geological features [269]; angular data of the Earths rotation was used to characterize the Chandler Wobble [109], while in physics, phase shift identification has applications to interferometry [85, 123] and in characterizing nano-scale surfaces [184]. In this Chapter, though, we focus, for concreteness and because of their intrinsic importance, on applications in EEG analysis and dynamical systems.

When neural activity is viewed on a large scale, such as in EEG recordings, desynchronization may manifest as a phase shift in the instantaneous phase of the signal (see §1.1). Among brain imaging techniques, the strength of EEG recording is that it has the high temporal resolution required resolve functional brain dynamics. For this reason, EEG recordings are often used in brain-computer interface (BCI) applications, where the goal is to provide an augmentative communication system which can interpret spontaneous brain activity [17, 298]. In such applications not only is it important to accurately identify events but it also must be done in real-time to facilitate timely feedback to the user. Analysis of phase behaviour in the alpha band (8-13 Hz) of EEG signals showed typical locking durations in the 150-450 ms range, while phase shift events occurred on a smaller time scale, lasting 45-67 ms [277].

Windowed methods of EEG analysis have been used to identify periods of phase synchronization using both Fourier transform [270] and Hilbert transform [190] definitions of phase (see §1.5), but these methods are unable to localize spontaneous phase shift events. Instead, the instantaneous phase difference derivative (see §2.2.2) may be used to chart the phase stability using the full temporal resolution of EEG recordings [93, 37, 277]. This approach has merit for the identification of phase shift events; however, a weakness of this approach is that identification has been done using a subjectively chosen threshold

value based on visual inspection of the data. A similar method using circular variance is used to identify global periods of phase synchronization, but this method also requires the subjective selection of a threshold parameter [60, 145].

Measures of phase synchronization are also used to quantify interactions within complex (chaotic) dynamic systems of oscillators [31]. One such system is the coupled Rössler attractor, where synchronization has been observed that is independent of the corresponding amplitude [209, 235, 222]. The experiments by Rosenblum [235] show that strong coupling results in complete synchronization; however, during weak coupling the system displayed dynamic transitions between phase locking and phase shifting behaviour. The existence of weak or intermediate coupling strengths is also a feature of systems which display self organized criticality (SOC), a class of dynamics which result in apparent power-law distributions [19]. These power-law distributions affect many physical aspects of the system, including the waiting time between events or inter-shift intervals (ISIs) [242]. Several studies have shown that functional brain networks display scale-free connectivity (i.e. power-law distributions in the degree of connectivity) [51, 260]. In addition, links between phase shift behaviour in EEG recordings and SOC were explored in [277], where power-law behaviour was observed in estimated power spectral densities of the PD function.

The goal of this chapter is to develop robust methods for phase shift identification from observables where we formulate the problem as a change-point problem. In Section 3.2 we describe instantaneous phase estimation, as well as the change-point methods which will be used. Section 3.3 contains simulations of simple oscillators, both with and without phase shift behaviour, which are used to assess the proposed methods in a controlled environment. In Section 3.4, methods are applied to data generated from a system of coupled Rössler attractors which have chaotic dynamics and spontaneous phase shift behaviour. In Section 3.5, the analysis is applied to EEG recordings where we identify phase shifts in the beta band of participants during a visual task.

3.1.1 Theory

Consider the instantaneous phase estimator in the case of an oscillator with additive i.i.d. noise and using a simple first-order filter. The exponentially weighted moving average (EWMA) is a tractable digital low-pass filter which is used to investigate properties of the estimator,

$$\mathbb{H}[x_t] = (1-\alpha)\sum_{i=0}^{t}\alpha^i x_{t-i}, \quad \alpha \in (0,1). \tag{3.1}$$

If the parameter is set to $\alpha = e^{2\pi f_c/T}$, then f_c is the -3 dB cutoff point for the filter (i.e. frequency components above f_c has been reduced by a factor of at least 10^{-3}).

Theorem 3.1. *Suppose x_t is a noisy oscillator sampled at T Hz with frequency f_0 Hz ($\omega = 2\pi f_0/T$) and constant phase ϕ i.e., $x_t = \sin(\omega t + \phi) + \epsilon_t$, with $E(\epsilon_t) = 0$ and $V(\epsilon_t) = \sigma^2$. Let $y_t = \mathbb{H}[x_t \sin\{-\omega t\}]$ and $\tilde{y}_t = \mathbb{H}[x_t \cos\{-\omega t\}]$ be the two components in the complex demodulation estimate of $\hat{\phi}_t$ (Eqn 1.2) with a EWMA(α) filter $\mathbb{H}[\cdot]$ given by (3.1). The expected values of y_t and $\tilde{y}_t$ are*

$$E(y_t) = \frac{\cos(\phi)}{2} + b(y_t), \quad E(\tilde{y}_t) = \frac{\sin(\phi)}{2} + b(\tilde{y}_t),$$

where $b(y_t)$ and $b(\tilde{y}_t)$ are the biases of y_t and $\tilde{y}_t$, respectively. We have that

$$\begin{aligned}
b(y_t) &= \frac{(1-\alpha)\big(\cos(2\omega t+\phi) + \alpha\cos(2\omega(t+1)+\phi)\big)}{2(1-2\alpha\cos(2\omega)+\alpha^2)} - \\
&\quad \frac{\alpha^{t+1}}{2}\left(\cos(\phi) - \frac{(1-\alpha)(\cos(\phi-2\omega) - \alpha\cos(\phi))}{(1-2\alpha\cos(2\omega)+\alpha^2)}\right), \\
b(\tilde{y}_t) &= \frac{(1-\alpha)\big(\sin(2\omega t+\phi) - \alpha\sin(2\omega(t+1)+\phi)\big)}{2(1-2\alpha\cos(2\omega)+\alpha^2)} - \\
&\quad \frac{\alpha^{t+1}}{2}\left(\sin(\phi) + \frac{(1-\alpha)(\sin(\phi-2\omega) - \alpha\sin(\phi))}{(1-2\alpha\cos(2\omega)+\alpha^2)}\right).
\end{aligned}$$

(b) The variance of Y_t is

$$\begin{aligned}
V(Y_t) &= \frac{\sigma^2}{2}\left(\frac{(1-\alpha)}{(1+\alpha)} - (1-\alpha)^2\left(\frac{\cos(2\omega t) - \alpha^2\cos(2\omega(t+1))}{1-2\alpha^2\cos(2\omega)+\alpha^4}\right) - \right. \\
&\quad \left. \alpha^{2t+2}\left(\frac{(1-\alpha)}{(1+\alpha)} - \frac{(1-\alpha)^2(\cos(2\omega)-\alpha^2}{1-2\alpha^2\cos(2\omega)+\alpha^4}\right)\right),
\end{aligned}$$

and for $\tilde{Y}_t$

$$V(\tilde{Y}_t) = \frac{\sigma^2}{2}\left(\frac{(1-\alpha)}{(1+\alpha)} + (1-\alpha)^2\left(\frac{\cos(2\omega t) - \alpha^2\cos(2\omega(t+1))}{1-2\alpha^2\cos(2\omega)+\alpha^4}\right) - \alpha^{2t+2}\left(\frac{(1-\alpha)}{(1+\alpha)} + \frac{(1-\alpha)^2(\cos(2\omega)-\alpha^2}{1-2\alpha^2\cos(2\omega)+\alpha^4}\right)\right).$$

(c) The covariance between Y_t *and* $\tilde{Y}_t$ *is given by*

$$Cov(Y_t, \tilde{Y}_t) = \frac{\sigma^2(1-\alpha)^2}{2}\left(\frac{\sin(2\omega t) - \alpha^2\sin(2\omega(t+1)) + \alpha^{2t+2}\sin(2\omega)}{1-2\alpha^2\cos(2\omega)+\alpha^4}\right).$$

Proof. By direct calculation (see Appendix D). □

The results of Theorem 3.1 allow us to understand the how the design of the filter (3.1), in particular the choice of the weight α and the distance from the boundary t, affects the estimate of the instantaneous phase. From these expressions we can see two distinct components in the bias; one purely oscillatory component and a 'boundary effect' which goes to zero for large t, since $\alpha^{t+1} \to 0$. An example of this behaviour with a 4th order Butterworth filter can be seen in Figure 1.2. The complex demodulation algorithm introduces a $2f_0$ Hz component into the signal and the oscillatory component represents the remains of this component after filtering. By increasing α in (3.1), the effectiveness of the filter is improved and the amplitude of the oscillatory component decreases, but this will lengthen the duration of the boundary effect.

The reliable identification of phase shift events relies on the magnitude of the shift being greater than the bias; thus the selection of a filter is a trade-off between the power to identify low magnitude shifts and the power to resolve shifts with small inter-shift-intervals (ISIs). For fixed α, and large values of t, the magnitude of the bias is bounded.

Corollary 3.1. *For a fixed* α *and large values of* t*, the bias of the instantaneous phase estimate is bounded by*

$$b(\hat{\phi}_t) \leq \tan^{-1}\left(\frac{(1+\alpha)}{\sqrt{2}(1-\alpha)-(1+\alpha)}\right) + O(\phi_t^2), \quad \alpha > \frac{\sqrt{2}-1}{\sqrt{2}+1}.$$

Proof. For large values of t, the boundary term goes to zero,

$$\lim_{t\to\infty} \alpha^t = 0.$$

To estimate the asymptotic bias in the estimator $\hat{\phi}_t$, we use a first order Taylor series expansion

$$E(f(y_t, \tilde{y}_t)) \approx f(E(y_t), E(\tilde{y}_t)).$$

This gives,

$$E\left(\tan^{-1}\left(\frac{\tilde{y}_t}{y_t}\right)\right) \approx \phi + \tan^{-1}\left(\frac{\cos(\phi)b(\tilde{y}_t) + \sin(\phi)b(y_t)}{1 + \cos(\phi)y(Y_t) + \sin(\phi)y(\tilde{Y}_t)}\right).$$

We can bound then bound the magnitude of the bias,

$$\begin{aligned} b(\hat{\phi}_t) &\approx \tan^{-1}\left(\frac{\cos(\phi)b(\tilde{y}_t) + \sin(\phi)b(y_t)}{1 + \cos(\phi)b(y_t) + \sin(\phi)b(\tilde{y}_t)}\right) + O(\phi_t^2) \\ &\leq \tan^{-1}\left(\frac{(1+\alpha)}{\sqrt{2}(1-\alpha) - (1+\alpha)}\right) + O(\phi_t^2), \quad \alpha > \frac{\sqrt{2}-1}{\sqrt{2}+1}. \end{aligned}$$

□

Although the theoretical results from this section are for a simple filter, they provide insight and intuition about the behaviour of more complicated filters, specifically the trade-off between size of persistent oscillatory bias and transient boundary effect (i.e. the portion of the bias which goes to zero for large t). The existence of a non-zero bias term means that there will be a lower bound on the magnitude of shifts (Δ_{min}) which can be reliably identified, and by reducing the noise levels we increase the range of identifiable events. Conversely, the boundary effect which occurs at the start of the recording will also be present at any phase shift event, and this will result in a minimum interval (ISI_{min}) such that two change point events can be accurately resolved. Decreasing the bandwidth in the complex demodulation algorithm (2δ) will decrease overall levels of noise, but lengthen the duration of transient boundary effects in the instantaneous phase estimate; the bandwidth represents a trade-off between statistical power and temporal resolution.

3.1.2 Change Point Analysis

Change-point analysis (CPA) is the identification of both the number and extent of time-dependent changes in model parameters. There are two primary approaches in CPA, parametric methods which require a specific model for the form of parameter changes (see [49] for a review), and nonparametric methods which do not make specific assumptions about the underlying distributions (see [39] for a review).

Early CPA methods use a cumulative summation (CUSUM) type statistic to test the hypothesis of no change point [211], a method which is often used in statistical process control [210]. Tractable exact or asymptotic distributions of CPA statistics can be determined for some simple parametric models [211, 252, 217] or nonparametric methods [26, 186, 217]. In other cases it is necessary to use computational techniques such as MCMC algorithms [50, 252] or bootstrapping techniques (see §3.2.2, [78, 6, 161, 162]) to determine the critical value for the test.

Methods have been developed specifically for change-point analysis (CPA) (or segmentation) in EEG recordings, using both parametric [225, 30] and nonparametric [40, 257] approaches; however, these amplitude based methods are not designed to capture changes in phase dynamics.

In CPA, a valid model for the anticipated changes (if such is known) will often provide the most powerful analysis. Here we are specifically interested in CPA on the mean of the instantaneous phase of the signal (see Section 1.1.1) which is a circular random variable [89]. There exist several methods of CPA for independent circular data; however, they give many false positive results in the absence of change points, due to the time-dependent bias and autocorrelations in our context. A parametric example is the test for a change in the mean of a Von Mises random variable [105, 115], a nonparametric approach is based on the rank CUSUM test [179].

3.2 Methods

3.2.1 Phase Shift Identification

The identification of phase shift events can be considered from two different perspectives, using the complete trajectory with fixed sample size, or real-time analysis which considers only information available prior to the current time. The latter occurs when it is necessary to have an immediate response to a shift event, such as in statistical process control or BCI; in this situation it is very important to identify events with computational efficiency and high temporal resolution. The current work focuses on the case of using a complete trajectory, while keeping in mind the potential for online identification.

To identify phase shift events, the estimated instantaneous phase is modelled as,

$$\hat{\phi}_t = \phi_t + b_t + \epsilon_t,$$

where ϕ_t is the true phase of the oscillator, b_t is a systematic time dependent bias and ϵ_t is random noise with $E(\epsilon_t) = 0$. We further assume that b_t is a periodic signal with an average value of zero,

$$\lim_{T \leftarrow \infty} \frac{1}{T} \sum_{t=1}^{T} b_t = 0$$

and that there exists a value $\tau \in \mathbb{Z}^+$ such that

$$Cov(\epsilon_t, \epsilon_{t+\tau}) = 0, \quad \text{and} \quad Cov(\epsilon_t, \epsilon_{t+k}) > 0 \quad for k < \tau$$

We test the hypothesis of no phase shift events ($H_0 : \phi_t = \phi_0$) against the alternative of a single phase shift event at an unknown time t_0 and with magnitude Δ ($H_a : \phi_t = \phi_0 + \Delta H(t - t_0)$, where H() is the Heaviside function). Two statistics are considered,

$$S_1 = \max_{2 \leq t \leq N-1} s_1(t), \tag{3.2}$$

$$S_2 = \max_{2 \leq t \leq N-1} s_2(t), \tag{3.3}$$

where

$$
\begin{aligned}
s_1(t) &= \left(\frac{N}{t(N-t)}\right)^{1/2} \sum_{i=1}^{t} \left(\hat{\phi}_i - \bar{\hat{\phi}}\right), \\
s_2(t) &= \frac{|\hat{\phi}_{t+1} - \hat{\phi}_{t-1}|}{2},
\end{aligned}
$$

and where $\hat{\phi}_t$ is the estimate of the instantaneous phase at time t and $\bar{\hat{\phi}}$ the temporal average. The former is a cumulative summation (CUSUM) type statistic which is favoured in change-point analysis (CPA), while the latter corresponds to a phase derivative (PD) method which is currently applied to EEG recordings.

The critical value of the hypothesis test (Φ_α), will be calculated using bootstrapping and approximation techniques (see §3.2.2 and §3.2.3). If the null hypothesis is rejected, a change point time $\hat{t}_0$ is estimated such that

$$
s_i(\hat{t}_0) = \max_t s_i(t), \quad i = 1, 2.
$$

Once $\hat{t}_0$ has been identified, the estimator can be iteratively applied to both halves of the sample to search for additional phase shift events. For the CUSUM estimator, the estimator is applied independently to the two halves of the sample, while for the PD estimator a pooled estimate of variance is used which includes all of the data. This process continues until either no more shifts are found, or there are less than the minimum amount of observations in the signal (N_{min}, estimated below) required to estimate the critical values using the bootstrap procedure. Note that the existence of a phase shift event at time t^* causes an additional boundary effect at the discontinuity of the instantaneous phase (i.e. an additional bias term which goes to 0 as $|t - t^*|$ goes to infinity). Such a boundary effect obscures the phase dynamics around the shift event. To ensure additional spurious phase shift events are not detected due to this boundary effect, an interval (t_L, t_U) about the estimated change-point must be excluded from further analyses.

For the CUSUM estimator, we use bootstrapped datasets to estimate the parameters N_{min}, t_L and t_U. Bootstrapped signals under the null hypothesis (no phase shift event) are

used to find the minimum value N_{min} which achieves the desired false positive rate,

$$P(S_1(x_{1:N_{min}}) > \Phi_\alpha) \leq \alpha.$$

Additionally, bootstrapped signals under the alternative hypothesis (single phase shift event) are used to estimate a value $ISI_{min}(\alpha)$, such that if

$$t_L = \hat{t}_0 - ISI_{min}(\alpha), \quad t_U = \hat{t}_0 + ISI_{min}(\alpha),$$

then the lower and upper subsamples $x_{1:t_L}$ and $x_{t_U:N}$ have the desired sampling distribution, that is,

$$P(S_1(x_{1:t_L}) > \Phi_\alpha) \leq \alpha, \quad P(S_1(x_{t_U:N}) > \Phi_\alpha) \leq \alpha.$$

The value ISI_{min} determines the temporal resolution for each method, i.e. the time scale for which multiple change point events can be accurately identified. The points t_L and t_U define the times when the amplitude of the boundary effect is reduced to the level of the persistent oscillatory errors in the signal. Once this has occurred the boundary effect can be properly accounted for by the bootstrap and approximation techniques and will not cause the method to identify additional false positive change points.

For the PD estimator, since this is an instantaneous estimator, we need not be so conservative in the exclusion of data around the change-point. In fact, it is only required to remove any points around $\hat{t}$ which are greater than the critical value,

$$t_L = \max_{t<\hat{t}_0}\{t|s_2(t) < \Phi_\alpha\}, \quad t_U = \min_{t>\hat{t}_0}\{t|s_2(t) < \Phi_\alpha\}.$$

3.2.2 Parametric Bootstrapping

To learn about the sampling distribution of estimators for a more complicated filter than considered in §3.1.1, a parametric bootstrapping procedure is used, [78]. Bootstrapping techniques are well developed for CPA in independent observations [6, 162]. To overcome the problem of dependencies, block bootstrapping techniques have been developed which preserve the temporal correlations in the bootstrapped datasets [165, 224, 126, 168]. These method have been successfully applied to CPA in problems with stationary dependencies [161].

Here we consider a simple 9 Hz ($f_0 = 9$) oscillator, sampled at T=250 Hz with constant unit amplitude and a fixed sequence of phase values ϕ_t. To account for the boundary effects in phase estimation, the first N_{burn} samples are removed from the analysis,

$$x_t = \sin\left(\frac{2\pi f_0 t}{T} + \phi_t\right), t = N_{burn} \ldots N.$$

In this parametric bootstrap, data are generated with i.i.d. additive noise ($\epsilon_t \sim N(0,1)$). Both the oscillator and noise term are normalized to unit power, and the signal-to-noise ratio (SNR) of the simulated observable is set by a weight parameter r,

$$x_t^{(b)} = \frac{r x_t}{||x_t||} + \frac{(1-r)\epsilon_t^{(b)}}{||\epsilon_t^{(b)}||}, \quad b = 1..B.$$

The relationship between r and the SNR is given by $SNR = 10\log_{10}\frac{r^2}{(1-r)^2}$ decibels.

For large values of N, the sampling distribution of the maximum value in a sequence of random variables converges to the generalized extreme value (GEV) family of distributions; this result has been proven for a broad class of dependent random variables, and does not require a stationarity assumption [102]. These results to not exactly cover the current context because there is also a time dependent bias; however, simulation exercises in Appendix D verify the convergence to a GEV distribution.

Additionally, Appendix D contains a simulation study which explores the behaviour of S_1 and S_2 in a simple parametric application. The results confirm the intuitive difference between a cumulative and instantaneous estimator. The S_1 (cumulative) estimator has greater power to identify phase shift events, especially with high noise or low effect size. Conversely, the S_2 (point-wise) estimator has better temporal resolution for resolving multiple shift events, i.e. it is able to identify multiple shifts with a smaller ISI.

3.2.3 Non-Parametric Methods

The above parametric procedures, where the noise is assumed i.i.d., provides a clear intuition about the range of shift magnitudes and SNRs where a shift can be reliably identified;

however, such methods are not appropriate for all situations. Neither the Rössler attractor nor EEG recordings can be accurately described by an ideal oscillator with i.i.d. noise. More often, signals from complex systems have oscillations which occur around, but not exactly on a fixed frequency. Other factors which can affect the detection of shift events are time dependent noise, or the existence of unrelated systematic features of the system such as oscillators at different frequencies.

To obtain an estimate of the sampling distributions without assuming some parametric form of the data, it is necessary to have some measure of the temporal dependence in estimated instantaneous phase. Here we use the first zero crossing of the autocorrelation function of $\hat{\phi}_t$ to estimate the length of the dependence (τ), although other methods may be used such as taking the local minimum of the time-lagged mutual information [92]. To estimate τ under the null hypothesis, the autocorrelation is estimated before unwrapping the instantaneous phase, preventing artificial inflation of $\hat{\tau}$ due to potential phase shift events in the signal. Additionally, due to the non-stationary nature of the signal, estimates of τ are combined from several different times in the signal. Here we take the arithmetic average of these estimates, though the median would also work.

Block Bootstrapping

To estimate the sampling distribution of the CUSUM estimator (S_1) under the null hypothesis, we employ the nonparametric block-permutation bootstrap technique. There are several different blocking techniques, notably the moving block [165] and circular bootstrap [224]. Here we use a non-overlapping block method [161], which is well studied in the change-point paradigm, the observed time series is partitioned into K intervals of length L,

$$x(k) = x_{1+L(k-1)} : x_{Lk}, \quad k = 1..K.$$

For a randomly generated permutation π of (1..K), the surrogate dataset is

$$x^{(b)} = [x(\pi(1)), x(\pi(2)), ..., x(\pi(K))].$$

Since the order of observations within each block remain unchanged, the surrogate datasets mimic the correlation structure of the underlying process. The value of L is chosen such that each window fully captures the autocorrelation structure of the signal, here we take a value of $L = 2\tau$.

Threshold Method

The method of block bootstrapping, is not applicable for the S_2 estimator. Creating a block permutation of the original data does not provide an accurate estimate of the critical value if the alternative hypothesis is true. This is because the nature of the estimator is instantaneous rather than cumulative, and all within block values of s_2 remain unchanged

We suggest an approximation to the distribution of S_2 which takes into account the variance and autocorrelation of the instantaneous phase. The phase differences are first centred and normalized to create a standardized sequence of N variables. To account for the dependence in the signal, the distribution of S_2 is approximated as the maximum of $K^* = 2\lfloor N/\tau \rfloor$ independent Gaussian random variables (where τ is defined as above, to be the first zero in the autocorrelation function). Critical values are based on the maximum of independent Gaussian random variables, although for large values of K^* it does not matter what specific distribution is used because the results will converge to a GEV distribution [102].

In the case of multiple phase shift events, rather than simply applying the method recursively, we suggest a pooled estimate of the variance which includes all parts of the signal which have not be identified as a phase shift event.

$$\hat{\sigma}^2_{pool} = \frac{\sum_{j=1}^{k}(N_j - 1)\hat{\sigma}^2_j}{\sum_{j=1}^{k}(N_j - 1)}, \tag{3.4}$$

$$\hat{\sigma}^2_j = \frac{1}{N_j - 1}\sum_{i=t_L(j)}^{t_U(j)}\left(s_1(j) - \bar{s_1}(t_L : t_U)\right)^2, \tag{3.5}$$

$$N_i = t_U(i) - t_L(i) + 1. \tag{3.6}$$

Thus the algorithm is:

1. Estimate the standard deviation of the phase locked values ($\hat{\sigma}^2_{pool}$)
2. Set the threshold to $\Phi_\alpha = \hat{\sigma}_{pool} z_{\left(\alpha^{K_i^*}\right)}$
3. Update the set of boundaries (t_L, t_U)
4. Test for phase shift events
    - If any new shift events are identified, update the set of phase shift events (t_L, t^*, t_U) and then go to step 1
    - Otherwise stop

As phase shift events are removed from the estimate, σ_{pool} will decrease, causing the critical value Φ_α to be monotonically decreasing. For this reason, the boundaries (t_L and t_U) must be updated at each iteration to account for the lowered threshold. Occasionally, lowering the threshold will cause two shift events to occur consecutively, in this situation it is not clear if there actually two events, so they should be merged into a single event rather than risk falsely declaring a spurious shift.

In addition to the increased temporal resolution, another benefit of the PD estimator is computational efficiency compared to the bootstrapping techniques presented for the CUSUM estimator. The current algorithm is described as a fixed sample size analysis; however, it could be adapted into an online (sequential) implementation by updating the values of $\hat{\sigma}_{pool}$ and K^* with each new observation and monitoring for values of s_2 which exceed the critical value.

3.3 Simple Oscillators

In this section the accuracy of the suggested methods is evaluated against the parametric alternative in the case of simple oscillators with i.i.d. noise. For each application, a fourth

order low-pass Butterworth filter is used to estimate the instantaneous phase because it has a maximally flat frequency response, i.e. it reduces the amplitude periodic variations (ripple) in the pass-band. To assess the methods, we use simulated oscillators which have a SNR of 0 dB and include $M = 20$ phase shift events. The shift magnitudes and ISIs are randomly drawn to obtain a robust comparison, and based on the results in Appendix D values of $\Delta_{min} = \pi/10 rad$ and $ISI_{min} = 1s$ are used,

$$\begin{aligned} \phi_t &= \sum_{i=0}^{M} \left(\sum_{j=0}^{i} \Delta_j \right) H(t - t_i), \\ \Delta_j &\sim Unif\left((-\pi, -\Delta_{min}] \cup [\Delta_{min}, \pi] \right), \\ t_i &= t_{i-1} + ISI_{min} + Exp(ISI_{min}). \end{aligned}$$

Each estimator is evaluated based on the rate of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN). The ability of each method to correctly detect change-point events are displayed for a range of significance level using receiver operating characteristic (ROC) curves, which plot TP against FP. The accuracy (ACC = (TP+TN) / (TP+FP+TN+FN)) provides a measure of goodness at specific significance levels, while the area under an ROC curve (AUROC) provides an overall measure of quality. The accuracy measure gives equal weight to TP and TN events, this is not necessarily optimal for every situation, other weightings may be considered depending on the application.

The resulting ROC curves from the change point analysis of twenty simulated oscillators (for a total of 400 shift events) are shown in Figure 3.1. Both the CUSUM and PD estimators have high power to identify phase shift events, while controlling the false positive rates. There is a cross-over of the ROC curves for the nonparametric estimators, which implies neither method is uniformly better than the other; the PD method provides the maximum accuracy, but the CUSUM method can provide increased power at the cost of a slightly higher FP rate. The parametric CUSUM method works best overall, and significantly better than the nonparametric CUSUM, while both the parametric and nonparametric PD methods give similar results. Table 3.1 shows the maximum ACC and AUROC values for each method, both nonparametric methods perform well, though the PD estimator is marginally better.

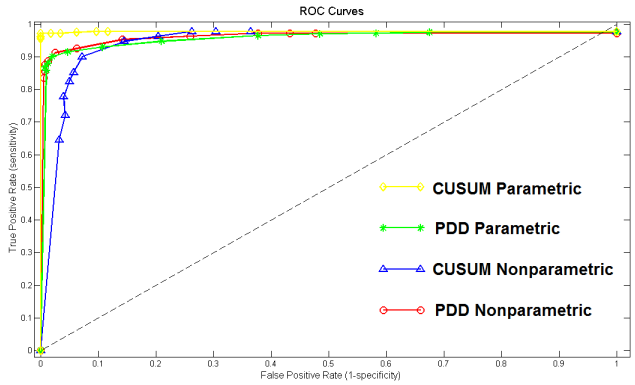

Figure 3.1: The ROC curves for both the parametric and nonparametric versions of the CUSUM and PD estimators. The parametric CUSUM estimator (yellow) performs the best, dominating each other curve. The parametric (green) and nonparametric (red) PD estimators perform similarly, they are both better than the nonparametric CUSUM (blue) for the low FP rates but then there is a cross-over and the nonparametric CUSUM performs better for high FP ranges. This is further addressed in the discussion, Section 3.5

Method	mACC	AUROC
CUSUM Parametric S_1	0.9863 ($\alpha = 0.1$)	0.9772
PD Parametric S_2	0.9400 ($\alpha = 0.01$)	0.9555
CUSUM Nonparametric S_1	0.9137 ($\alpha = 0.1$)	0.9438
PD Nonparametric S_2	0.9438 ($\alpha = 0.03$)	0.9610

Table 3.1: A table of the maximum accuracy (mACC) and area under the curve (AUROC) for the simple oscillators. Both nonparametric measures provide comparable results, although the instantaneous estimator (S_2) has both the highest mACC and AUROC.

3.4 Rössler Attractor

In this section, we apply the proposed methods to a system of coupled Rössler attractors. This application represents an intermediate level of difficulty between the i.i.d. oscillators and EEG application. The Rössler attractors chaotic dynamics are not strictly periodic and significant power leaks into nearby frequency bands, and they are more sophisticated than the simple oscillator and similar to the spectral properties of EEG signals; however, unlike an EEG application the truth of shift events in the Rössler attractor are still available to calibrate methods.

Coupled Rössler attractors (see §1.1.2) are simulated with parameter values $a = 0.15$, $b = 0.2$, $c = 10$. The frequency parameters were $\omega_1 = 2\pi f_0 + \delta\omega$ and $\omega_2 = 2\pi f_0 - \delta\omega$, where the attractors have average frequency of $f_0 = 9Hz$ and a frequency mismatch of $\delta\omega = 0.675$. For this set of frequency parameters, the phase shift dynamics observed in [235] are recreated by setting the coupling parameter to $C = 0.12$. Trajectories were generated at a rate of 10 kHz and then down-sampled to 250 Hz. Estimated power spectral densities are shown in Figure 3.2, the attractor shows a distinct peak at approximately 9.25 Hz.

To estimate the autocorrelation which remains in the instantaneous phase, we increase the coupling value to C=0.5 to generate signals with strong synchronization and no shift events. The ACF of the instantaneous phase difference for the strongly coupled oscillators is shown in Figure 3.2. This first zero crossing of the ACF function occurs at $\tau = 1183$, or approximately 4.5 seconds.

3.4.1 Rössler Results

To assess the ability of each method to identify phase shift events in the weakly coupled Rössler attractors, we generated fifty datasets with random initial conditions and a length of 10 minutes. Phase shift events were manually marked using the Poincare definition of

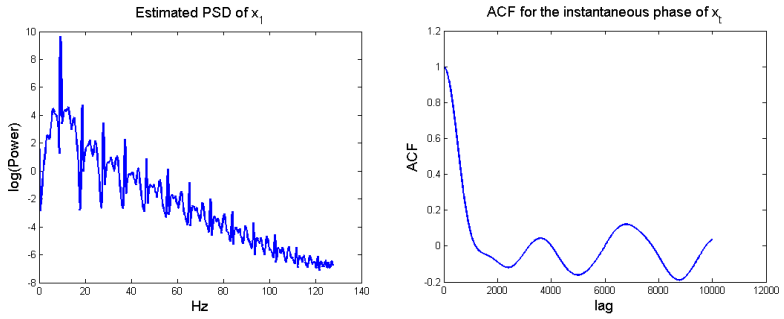

Figure 3.2: Left: Estimated log PSD of x_1 for the coupled attractors (densities based on x_2, $y_{1/2}$ were nearly identical). There is a distinct peak at approximately 9.25 Hz. The signal is not strictly periodic and the spectral components 'leak' into other bands. Notably, there is a periodic beat frequency due to the frequency mismatch ($d\omega$) of the oscillators. Right: Estimated ACF function of the instantaneous phase of x_t during a period of phase synchronization (no shift events). The first zero crossing of the ACF can be used to determine the appropriate value of L for the block bootstrapping algorithm. Here the crossing occurs at a lag of 1183, or approximately 4.5 seconds.

phase. There were 139 total phase shift events for an average of 2.78 per dataset.

Using the observable $h = x_t$, shift events are estimated using both nonparametric methods. An ROC curve of the results (see Figure 3.3) shows that the CUSUM estimator outperforms the PD estimator. There are several potential reasons why the CUSUM estimator outperforms the PD in this situation, in contrast to independent oscillators; (1) longer average ISI's, (2) uni-directed phase shifts, and (3) temporally correlated signal features. The PD estimator eventually crossed the CUSUM estimator in the ROC curve, this is due to the few shift events which occur with an ISI that is too small to be resolved by the CUSUM estimator. The maximum accuracy (mACC) and area under the curve (AUROC) for these methods are shown in Table 3.2, here we see that both estimators perform well but the CUSUM estimator performs best.

To investigate potential power law behaviour in the distribution of ISIs, we generate one thousand twenty-minute datasets with random initial conditions. Applying the CUSUM

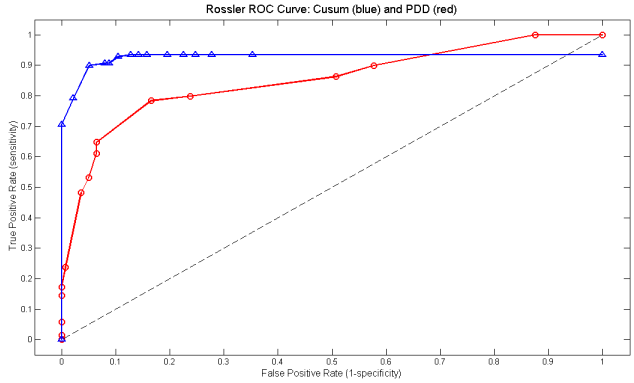

Figure 3.3: Plot of the ROC curve of the nonparametric CUSUM (blue) and PD (red) phase shift estimators, applied to simulated data from coupled Rössler attractors. For this application, the CUSUM estimator outperforms the PD estimator in both mACC and AUROC. The intervals between shift events from the attractors are long enough that the CUSUM method identifies over 90% of them. The PD is able to identify all shift events, but does so with high false positive rate, due to the temporally properties of the attractor dynamics.

Method	mACC	AUROC
Nonparametric CUSUM S_1	0.8225 ($\alpha = 0.05$)	0.9238
Nonparametric PD S_2	0.6187 ($\alpha = 0.1$)	0.8504

Table 3.2: A table of the maximum accuracy (mACC) and area under the curve (AUROC) measures of the ROC plot for coupled Rössler attractors. For this application the CUSUM estimator outperforms the PD estimator.

estimator with $\alpha = 0.05$ found a total of 5661 phase shift events, with an average ISI of $\mu = 2.72$ minutes ($\sigma = 1.66$). Histograms of the observed ISIs are shown in Figure 3.4, on standard and log-log scales; the log-log histogram clearly shows an asymptotic power-law behaviour with an estimated slope of $q = -3.92$.

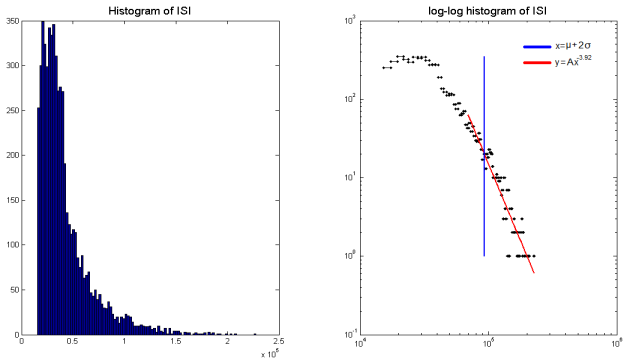

Figure 3.4: Left: Histogram of the observed ISIs in the weakly coupled Rössler attractors, as estimated by the CUSUM estimator. Right: corresponding log-log scale histogram, there is a linear relationship in the tail of the distribution, with a scaling exponent of 3.92.

3.5 EEG Phase Shift

This section applies the methodology to EEG recordings. These signals are comprised of many different components, and as in the case of the Rössler attractors, there is power in multiple spectral bands. In this real world problem there is no information on the truth of phase shift events to calibrate the algorithms, thus the utility of the methods is assessed by their ability generate neurologically plausible results. Relationships between shift events and external visual stimuli are explored, with emphasis on the scalp regions which are associated with such a visual task. Additionally, the distribution of ISI's is explored, looking for evidence of the asymptotic power-law distribution which are hypothesised to exist [260].

For this application we focus specifically on recordings during the visual vigilance task from the EEG dataset (see §1.1.3). The beta band (13-30Hz) of EEG recordings is often associated with sensorimotor activity [218], such as in the flanker task in our data. A recent attempt at a unifying hypothesis of the functional role of beta band oscillations suggests that it is responsible for maintenance of sensorimotor or cognitive state [82]. Spectral power analyses of the beta-band has been previously employed for classification in BCI

applications [17]. Here we focus specifically on the values $\omega = 16.5$ and $\Delta\omega = 3.5$, corresponding to the (13-20Hz) lower beta band.

To estimate the length of the autocorrelations in the EEG application, we first divide the recordings into 4 second segments. For each segment and each pair of channels, we estimate first zero-crossing (τ) in the ACF of the wrapped instantaneous phase. The value of K is calculated based on the average of all the estimated values of τ, this results in a value of $K = 85$.

For each pair of signals, we apply the PD identification algorithm to the instantaneous phase difference of the pairs, to identify spontaneous desynchronizations. In total there were 4 690 214 phase shift events across all 18 participants and $16 \times 15 = 120$ pairs of electrodes, with an average ISI of 262 ms (standard deviation 226 ms) or 3.8 shifts per second. We also investigate power-law behaviour in the distribution of ISI of beta band phase shift events. Standard and log-log scale histograms of the distribution of ISIs are shown in Figure 3.5. There appears to be asymptotic power-law behaviour in the tail of the distribution with a scaling exponent of $q = 6.26$

We further consider the relationship between the occurrence of phase shift events and the task stimuli. For each pair of signals (i,j), and each shift event (k), we record the amount of time since the most recent stimulus event. If there is no relationship between the shift and stimuli, then we expect that the times will be uniformly distributed. To test this hypothesis, we group the variable into 10 equal sized bins, between 0-500 ms and apply a χ^2 test for uniformity. Results from this analysis are summarized in Figure 3.6; there are three pairs of sites (Fp1-F7, T3-O1, T6-O1) which are significant at the $\alpha = 0.05$ level, including a Bonferroni correction ($p < \alpha/120$) and an additional six pairs (Fp1-T3, Fp1-O1, Fp2-O1, T3-T4, T5-O1, F8-P3) which are significant with a less conservative correction ($p < 0.05/30$). As expected for our visual attention task, many of these pairs involve the occipital (O1,O2) sites over the visual cortex and prefrontal sites (Fp1, Fp2), which are associated with attention.

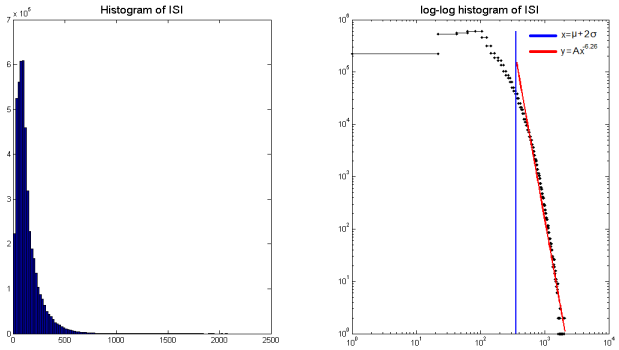

Figure 3.5: Left: Histogram of the observed ISIs in the beta band activity of EEG recordings during a visual vigilance task, as estimated by the PD estimator. Right: The corresponding log-log scale histogram, there is a linear relationship in the tail of the distribution, with a scaling exponent of 6.26.

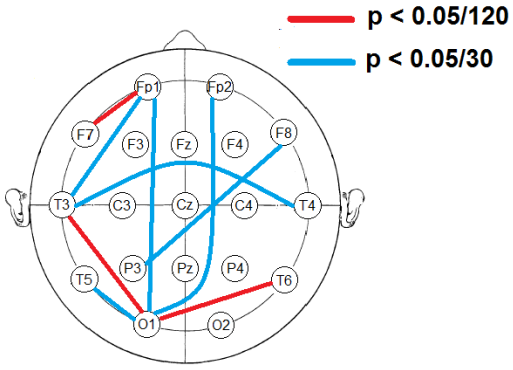

Figure 3.6: The standard 10-20 system introduced in [147]. Results of the tests for uniformity of phase shift events between visual stimuli. Red lines represent significance with a full Bonferroni corrected p-value ($p < 0.05/120$) while blue lines represent significance at a less conservative level ($p < 0.05/30$). The region most commonly related to the stimuli is the left occipital region O1, directly above the visual cortex. Additionally, the frontal sites Fp1 and Fp2, commonly associated with attention, are also related to the stimuli.

Discussion

In some EEG applications such as brain computer interfacing (BCI), it is desirable to not only identify events with high temporal resolution, but also in real-time so that feedback can be provided immediately. This allows the participant to work towards controlling some external device, such as a computer or prosthetic. The instantaneous nature of the PD estimator, as well as its computational efficiency, make it easily modified to perform in real-time. Conversely, there is no obvious analog for the CUSUM estimator, which requires access to the entire signal, as well as time consuming bootstrap procedures, and as such is better suited to post-hoc analysis.

Regarding the stability of the stability of the proposed methods, in the case of multiple shift events it is possible that small changes in parameters may have a large effect on results of the CUSUM estimator. If the first change-point is not identified, then all other potential shifts are not identified. This is not the case in with the PD estimator, the instantaneous nature of s_2 means that when a change point is found it does not affect the value of s_2 at other time points, resulting in a more stable algorithm. This idea can be seen in the results shown in Figure 3.1, where increasing the false positive rate (α) has a larger affect on the CUSUM nonparametric CUSUM estimator and the ROC curve for that estimator crosses over the more stable ROC curves for the PD estimators.

There are often many sources of noise in measuring observable time series from complex systems; systematic effects, additional oscillating components, temporally correlated noise, spatially correlated noise (i.e. source mixing in EEG due to volume conduction), frequency misspecification or narrow-band signals (non-fixed frequency). An investigation into the effect of such features in controlled environment may provide insight as to how to formulate a more robust solution to the shift identification problem.

In summary, this Chapter investigate the identification of phase shift events from observed time series, a problem which occurs in complex systems across many fields. With respect to the identification of phase shift events from EEG recordings, this Chapter pro-

vides an objective method for shift identification which may be applied to PSGC or PR analysis, providing greater confidence in the results of such analyses, as well as reproducibility across labs.

In Chapter 4, I perform a simulation study to assess the proposed estimators in the context of PSGC analysis. Concurrently, I will assess the effect of volume conduction on the PSGC measure to determine. The PSGC application presented in Chapter 2 is also repeated with the updated identification procedure, and taken with the results of the simulation study, conclusions are drawn on the connectivity patterns of different tasks.

Chapter 4

Volume Conduction

In Chapter 2 we introduced a novel measure of connectivity analysis for use with EEG recordings, based on spatio-temporal patterns of phase shift events. In this chapter a simulation study is undertaken to verify that the proposed Phase Shift Granger Causality (PSGC) method is able to identify connectivity patterns in situations similar to electroencephalogram (EEG) recordings. In particular that it can deal with issues such as high levels of noise and linear source mixing.

It is shown that, in fact, the PSGC method is able to correctly identify both the existence and direction of information transfer, and that the existence of spatio-temporal noise serves to reduce the spread of shift identification due to volume conduction (VC). When dealing with a linear combination of oscillators, the instantaneous phase of a sum is an amplitude weighted function of the phase from the individual components; as the amount of noise increases, the source of the phase shift is less likely to be the dominant component of the signal as the distance from the source increases. An example of this effect is shown in Figure 4.1, where the spread of phase shift events is compared for two different levels of the biological signal-to-noise ratio (bSNR, a spatio-temporal noise rather than independent measurement error). The spread of shift events is larger in the high bSNR case than in the low bSNR, showing the higher levels of spatio-temporal noise attenuate the spread of shift event. Full details on how that spread of activity was calculate can be found in §4.3 and

the specific details of this example in Appendix D.

Additionally, the simulation study is used to assess methods of phase shift identification introduced in Chapter 3. The subjective method used in Chapter 2 is compared with other objective methods of determine a threshold (critical) value when using the phase derivative (PD) estimator. Based on the results of the simulation study, the PSGC analysis is reapplied to the EEG data (see §1.1.3) with an objective threshold technique. From these results we have increased confidence in our conclusions that the PSGC is able to clearly distinguish between the resting task and the active tasks; the latter have higher rates of connectivity overall, and specifically more long range connections.

In this chapter, my contributions to the field is a quantification of the effect of VC and reference electrode on the identification of phase shift events, and also the estimation of connectivity using the PSGC measure introduced in Chapter 2.

4.1 Introduction

At any point in time, EEG signals can be modelled as a linear combination of the activity from each contributing electrical source. The conductive properties of the head instantly create attenuated versions of source activity at every scalp electrode, a process known as volume conduction (see §1.5.1, §1.6.1, Appendix A). In studies of EEG coherence, VC creates artificially inflated coherence estimates between distant electrodes [205]. When considering the effect of VC on connectivity estimation, there are two different situations; the existence of spurious connectivity due to a single neural source projecting to multiple scalp sites, and mislocalizing the source of legitimate neural connectivity.

There are three approaches which are used to account for the effects of VC in EEG analysis; experiment designs (contrasts) [277], connectivity measures which are designed to minimize the effect of VC [201], or by employing source reconstruction techniques [249]. In this Chapter I focus on the effect of VC on estimates of connectivity at the scalp, while

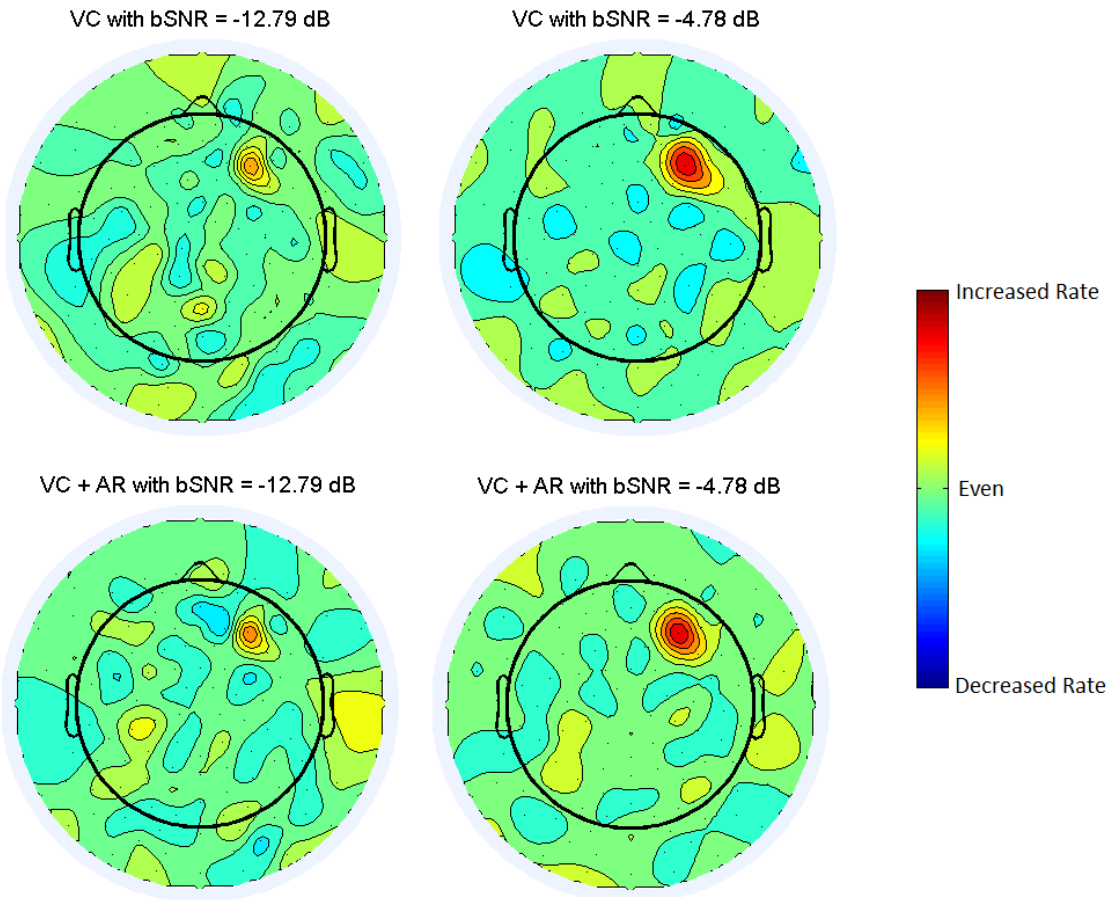

Figure 4.1: Effect of volume conduction (VC - Top) and volume conduction + average reference (VC + AR - Bottom) on phase shift identification. Topographies show the average number of shift events due to a generator at F4 and spatio-temporal noise with bSNR levels of -12.79 dB ($r_b = 0.05$ - Left) and -4.79 dB ($r_b = 0.25$ - Right). In the -12.79 dB condition the effect of VC and average reference is highly attenuated by the biological noise; the spread is only to the nearest neighbours of the high density 128 electrode net. In the -4.78 dB condition there is increased spread due to volume conduction but still less than the inter-electrode distance in the 10-20 montage. There is also a mild increase in connectivity due to the average reference at this level of bSNR. Topographies created using EEGLAB [71].

the use of source reconstruction techniques is explored in Chapter 5.

There are two properties that are exploited to reduced the effect of VC by experiment design, that the effect of VC is determined by the conductive properties of the particular head being measured, and that the effect decreases quadratically with distance (see Eqn A.2). One approach is to only look at differences in connectivity across tasks, keeping the effect of VC static. A common misconception is that differences between tasks must be the result of change in the underlying connectivity. Rather, the effect of VC on EEG connectivity is modulated by noise levels in the recording [249], so that what appears to be differences in connectivity between tasks may be varying noise profiles instead of changes in connectivity. Another experimental design only estimates connectivity from groups of electrodes which are aligned in a row [277]. In this scenario, if VC is affecting connectivity between distant electrodes than it must also affect intermediate pairs, and the absence of connectivity in the intermediate pairs implies that the longer effect is not due to VC.

Measures of connectivity are explored [201, 202, 264] which are (theoretically) immune to, or at least are less affected by VC. Note, for example that the spread of activity due to VC is essentially instantaneous (see Appendix A), so Granger causality analysis should not be affected; however, connectivity estimates may still be distorted due to 'weak' asymmetries such as a difference in the signal-to-noise ratio [135]. Imaginary coherence is a method which omits the real part of coherence, which theoretically contains all effects of VC [201]; however, it is unknown how much true information is lost by this process. Related methods which works on a similar assumption is phase slope index [202] and phase lag index [264], in which in-phase components do not contribute to connectivity.

In this chapter, I present a proof-of-concept simulation study to determine the effectiveness of PSGC to estimate connectivity from high-noise, non-averaged EEG recordings. This will include an investigation into the spread of phase shift events at the scalp due to VC and an average reference, as well as to assess the overall effect of VC on the PSGC analysis. The identification of phase shift events will be performed with several different methods of shift identification, to consider the effect of different threshold selection pro-

cedures. Finally, based on the conclusions of the simulation study, I re-analyze the EEG study from Chapter 2 and draw conclusions on connectivity results.

4.2 Methods

In this chapter, I apply the PSGC method as described in Chapter 2. For the identification of phase shift events, based on the results of Chapter 3, the PD method of shift identification is used for its superior temporal resolution.

4.2.1 Phase Shift Identification

Three different methods of shift identification are considered; the PD estimator from Chapter 3 and two threshold methods from Chapter 2, one threshold chosen subjectively and the other using an objective estimate modified from the PD estimator,

$$c = \hat{\sigma} z_{(1-\alpha)}. \tag{4.1}$$

The threshold is estimated as the $(1-\alpha)$ quantile of a Gaussian distribution, this is essentially the critical value estimated in Chapter 3 using $K^* = \frac{1}{2}$ ($\tau = \frac{N}{2}$) instead of estimating it from the data. Using this estimate will result in a less conservative approach (lower threshold) than the PD estimator, by not taking into account the length of the signal when determining the critical value.

Since this estimator does not take into account the length of the signal, a common threshold value can be used for the entire signal. Phase shift events in the signal will artificially increase the estimated variance of the instantaneous phase derivative, so as phase shift events are identified those points should be removed from the variance estimate. If the maximum value of $|\phi'_t|$ exceeds the threshold value, each point above the threshold is removed from the set of phase locked values, and then the sample variance (and threshold) are recalculated using only phase locked values. The process is then recursively applied until the threshold value has converged to a fixed value. Occasionally the threshold may

oscillate between a small number potential threshold values, to avoid this situation, if ever the threshold will increase the algorithm terminates instead. Starting under the assumption that the entire signal is phase locked, the algorithm for calculating the threshold is:

(1) Estimate the standard deviation of the phase locked values ($\hat{\sigma}^2$)

(2) Set the threshold to $c = \hat{\sigma} z_\alpha$

(3) Update the set of phase locked values, $A = \{t, |\phi'_t| < c\}$

(4) Update the estimated standard deviation ($\hat{\sigma}^2_A$) and threshold value ($c^* = \hat{\sigma}_A z_\alpha$)

(5) If $c^* < c$ then set $c = c^*$ and go to (3), otherwise stop

Once the threshold has converged, a phase shift event is identified at the value of t which maximizes $|\phi'_t|$ for each interval of time where it is above the threshold.

4.3 Proof of Concept

To verify that the PSGC method is appropriate for EEG analysis, we performed simulations of simple oscillators with coupled phase shift behaviour. The data were then transformed with a realistic lead field matrix to emulate the linear mixing characteristics of EEG. Simulations were performed for several levels of coupling strength, biological and measurement noise, to identify the range of behaviours for which the PSGC method reliably identifies connectivity.

4.3.1 Simulated Data

Simulated data were generated from two oscillators located under the F4 and O2 scalp regions at a depth of 15% below the scalp (corresponding to the surface of the brain [239]), each with a 9 Hz component of interest. The phase of each oscillator was generated as

$$\phi_{i,t} = \phi_{i,t-1} + \epsilon_{i,t}\eta_{i,t}, \quad i = \{F4, O2\}. \tag{4.2}$$

where $\epsilon_{i,t}$ are binary variables which determine whether or not a phase shift occurs and $\eta_{i,t}$ are independent random variables drawn from a $N(3\pi/5, \pi/5)$ distribution, which represent the phase shift magnitude. Each generator had a fixed probability of a shift event occurring (γ_i), unless the generator was in a brief refractory period which occurred after each shift. The refractory period was a fixed 500 ms (approximately 5 cycles of 9 Hz) during which there was no potential for a phase shift event. To include information flow into the model, coupling parameters were used to increase the rate of phase shift events in the receiving generator for a period after each shift at a driving generator. Defining $\bar{\epsilon}_{i,t}$ as the number of shifts at generator i which have occurred in the previous 500 ms, the probability of a phase shift is given by

$$P(\epsilon_{i,t} = 1) = \begin{cases} \gamma_i \quad if \quad \bar{\epsilon}_{i,t} = 0, \bar{\epsilon}_{\setminus i,t} = 0, \\ \gamma_i + C_i \quad if \quad \bar{\epsilon}_{i,t} = 0, \bar{\epsilon}_{\setminus i,t} = 1, \\ 0 \quad otherwise, \end{cases} \tag{4.3}$$

where $\setminus i$ represents the driving generator. For these simulations the base probability of shifting was set to $\gamma_i = 0.0024$, resulting in approximately one shift every 0.83 seconds (approximately the rate of shift events observed in the EEG data). Here we were primarily interested in the ability to identify an asymmetric coupling from the generator under F4 to the generator under O2 ($C_{F4} = 0,\ C_{O2} > 0$).

Once instantaneous phase values were randomly generated, they were then converted to oscillators with unit amplitude, frequency $f = 9$ and sampling rate $T = 500$ Hz,

$$x_{i,t} = \sin\left(2\pi f t/T + \phi_{i,t}\right).$$

The signals ($X_t = [x_{F4,t}, x_{O2,t}]$) from the generators were then linearly mixed to 'scalp electrodes' using a realistic lead field matrix (A_x, see Appendix A) with the generators oriented to be perpendicular to the scalp [192]. Sample scalp topographies for the simulated data are shown in Figure 4.2,

$$S_t = A_x X_t.$$

Following closely with [135], different types of noise were included in the simulations; however, here we defined separate parameters to control levels of biological signal-to-noise ratio

(bSNR: $r_b \in [0, 1]$) and measurement signal-to-noise ratio (SNR: $r_m \in [0, 1]$). Independent measurement noise (E_t) was simulated from a N(0,1) distribution for each sensor. Additionally, several spatiotemporal sources of biological noise ($n_{i,t}$) were generated which represent other biological signals such as unrelated brain activity. The biological noise sources were generated from stationary AR(10) processes, with coefficients independently drawn from a N(0, 0.01) distribution. Spatial locations for the biological noise ($N_{biol,t} = [n_{1,t}, n_{2,t}, ...n_{10,t}]$) were uniformly drawn from each hemisphere (five on the right, five on the left) and placed below the scalp with depths selected uniformly from (15, 40)% of the head radius. The biological noise was then linearly mixed to the same scalp locations using a matrix A_n which was calculated using to the same method as A_x,

$$N_t = A_n N_{biol,t}.$$

The signal and biological noise were scaled to have the specified contributions to the overall biological signal,

$$B_t = \frac{r_b S_t}{||vec(S_t)||} + \frac{(1 - r_b)N_t}{||vec(N_t)||}. \tag{4.4}$$

The overall biological signal B_t and measurement noise E_t were both normalized and the SNR was set in the observed data Y_t using a weighted combination of the two,

$$Y_t = \frac{r_m B_t}{||vec(B_t)||} + \frac{(1 - r_m)E_t}{||vec(E_t)||}. \tag{4.5}$$

All sources were mixed to 128 scalp sites referenced to Cz and then rereferenced to the average, to be consistent with the preprocessing applied in the application to real EEG data. Data were then reduced to eight channels at right hemisphere locations (Fp2, F4, C4, P4, O2, F8, T4, T6) for the PSGC analysis.

4.3.2 Application Details

In this section we apply the PSGC method in a proof-of-concept simulation study (Section 4.3) as well as to real EEG recordings of resting and vigilance tasks (Section 4.4). Here we describe the specific details of the implementation of the PSGC method for both the real

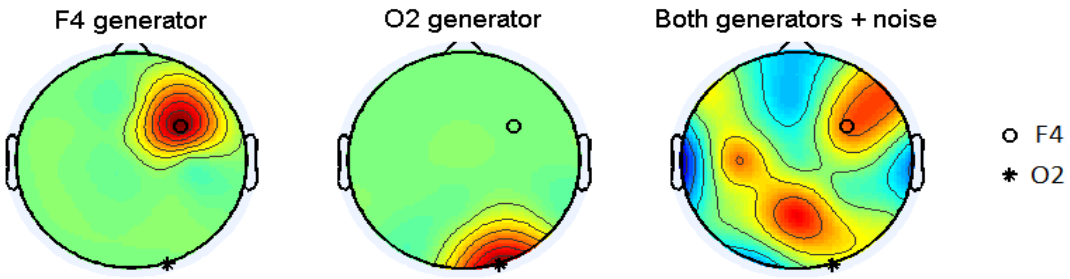

Figure 4.2: Scalp topographies of the two generators (Left: F4, Centre: O2) and then both generators with the inclusion of spatial/temporal and measurement noise. Although the simulation study is performed solely on the right hemisphere sites, a full montage of electrodes with measurement noise is used to generate these topographies. All topographies were taken at time $t = 1$, a SNR of 0 dB ($r_m = 0.5$), a bSNR of -4.8 dB ($r_b = 0.25$) and were generated using EEGLAB [71].

and simulated applications.

Mirroring the analysis in the EEG application the frequency band of interest for these analyses was 8-10 Hz. The instantaneous phase of each channel was calculated by shifting the centre of the band of interest to 0 Hz and then putting the result through a 1 Hz low pass Butterworth filter. Calculation of the threshold value was performed using each of the methods in §4.2.1; the subjective threshold is set at $c = 0.02$, and for other methods a value of $\alpha = 0.05$ is used. A value of $W = 100$ ms was used for the length of the history windows, this value of W is approximately the length of a complete cycle at the frequency band of interest. For each simulated dataset and each channel, model selection was performed to select the optimal order M for the log-linear parametrization. The optimal value was determined by calculating the Akaike information criterion [4] for a range of values of M (1 to 6) and selecting the one which minimized the criteria.

Results from statistical analyses were corrected for multiple comparisons using the Holm-Bonferroni procedure [141] to give a family-wise error rate of $\alpha = 0.05$ per hemisphere. For individual networks estimated using the likelihood ratio test (LRT, see Eqn 2.3) corrections are applied based on (8 x 7) = 56 comparisons per hemisphere and (8 x 7)/2 = 28 multiple comparisons for the net Granger causality (netGC, see Eqn 2.5) analysis.

4.3.3 Simulation Results

Simulations were performed with three kinds of coupling (none, directed, feedback) over a range of bSNR values (see §4.5 for a discussion of the plausible range of bSNR values for EEG data). Preliminary results suggested that independent measurement noise does not contribute strongly to PSGC results so the SNR was set to a fixed 0 dB ($r_m = 0.5$) for all simulations. For each set of parameters, 18 datasets were generated (to match the 18 participants in the application to real data). Simulated datasets were 5 minutes in length for the no-coupling and directed-coupling scenarios. Simulations of the feedback scenario were 10 minutes in length, as it was necessary to increase the power of the PSGC measure to identify this more complex dynamic.

The results of the PSGC analysis for all possible pairs of connections is summarized graphically as connectivity matrices for several sets of parameters. Each cell of the matrix represents the Granger causality from one signal to another. The i,j element of the matrix represents PSGC from N_i to N_j. The shade of the cells represent the number of simulations with significant connectivity in the specified pathway. The i,j element is white if causality exists in every participant, black if there is causality in zero participants and shades of grey represent intermediate results. The diagonal entries are crossed out to emphasize that there is no analysis on these locations. We further explore the relationship between bSNR and PSGC by looking at the connectivity rates for two specific connections of interest ($F4 \rightarrow O2$ and $O2 \rightarrow F4$). Plots of the connectivity rates as a function of bSNR are explored for simulations with no coupling ($C_{F4} = C_{O2} = 0$), directed coupling ($C_{F4} = 0$, $C_{O2} = 0.015$) and feedback ($C_{F4} = C_{O2} = 0.015$).

Subjective Threshold

Panels A and B of Figure 4.3 show the results of the subjective threshold with directed coupling ($C_{F4} = 0$, $C_{O2} = 0.015$), and bSNRs of -12.8 db ($r_b = 0.05$) and -4.8 dB ($r_b = 0.25$). The true connection ($F4 \rightarrow O2$) is the only connection identified in multiple simulations, and the power to identify this connection increases with bSNR. At both bSNR levels there are minimal false positive connections (never in more than 1 of the 18 simulations for a given pathway). Panels C and D show the results of simulations with no coupling ($C_{F4} = C_{O2} = 0$), there are minimal connections identified; for bSNR levels of -4.8 dB and lower, the number of spurious connections which are identified across the four conditions (10, or 13%) is slightly larger than the number of false positives we expect across many simulations (3.6 or 5%).

In the left panel of Figure 4.4 we see that the PSGC measure has high power to identify the true directed connection for a range of bSNR values, power of up to 50% to identify the feedback connection, and in the uncoupled case we see no false positives for bSNR values up to -4.8 dB ($r_b = 0.25$). In the right panel, we see the same power to identify feedback connectivity, but only a single spurious connection in both the no-coupling and directed-coupling cases for bSNR levels of -4.8 dB ($r_b = 0.25$) or lower. The reduced power to identify feedback connectivity is due to the inseparability of the signals in this paradigm, two signals become entangled and it becomes difficult to isolate the causal effects. This is an acknowledged limitation in Granger causal analysis [116], other measures of causality may be necessary for this situation [266].

Objective Threshold

Panels A and B of Figure 4.5 show the results of the objective threshold with directed coupling ($C_{F4} = 0$, $C_{O2} = 0.015$), and bSNRs of -12.8 db ($r_b = 0.05$) and -4.8 dB (and $r_b = 0.25$). Similar to the results of the subjective threshold, the true connection ($F4 \rightarrow O2$) is the only connection identified in multiple simulations, with the power to

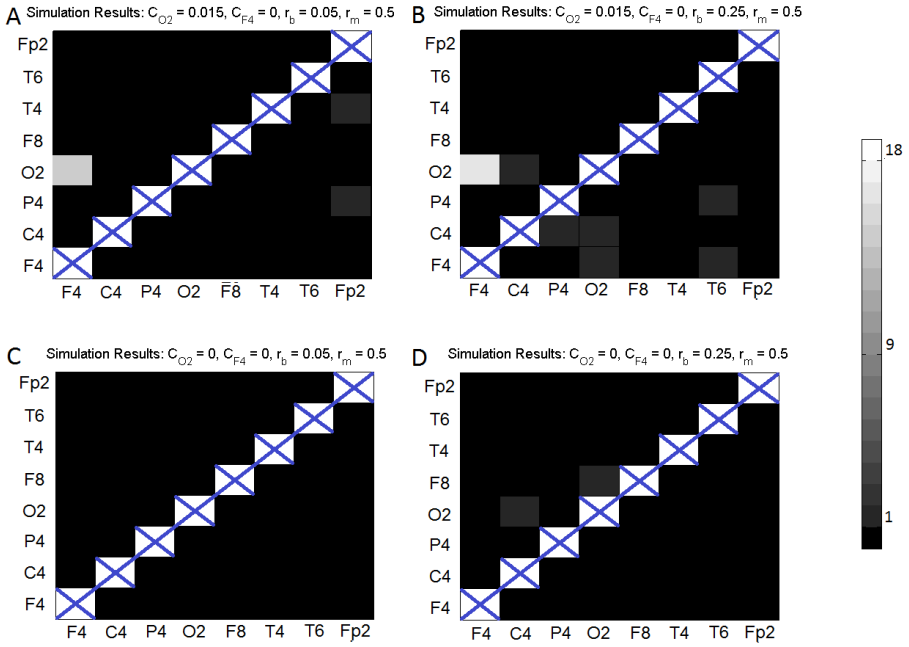

Figure 4.3: Each cell (i,j) of the matrix represents the Granger causality from from N_i to N_j. The shade of the cells (see colourbar) represent the number of simulations which found significant connectivity in the specified pathway. Panels A and B show the results of simulations with asymmetric coupling ($C_{F4} = 0$, $C_{O2} = 0.015$); at bSNRs -4.8 dB (and $r_b = 0.25$) the true connectivity is identified in 16 of 18 simulations ($F4 \rightarrow O2$) and there is only sparse connectivity otherwise (never in more than 1 of the 18 simulations for a given pathway). Panels C and D show the results of simulations with no coupling ($C_{F4} = C_{O2} = 0$), where there are few spurious connections, indicating minimal effect of VC.

identify increasing concurrently with bSNR. Panels C and D show the results of simulations with no coupling ($C_{F4} = C_{O2} = 0$), there are minimal connections identified; overall, the number of spurious connections (8, 11%) is lower than in the subjective threshold case (10, 13%).

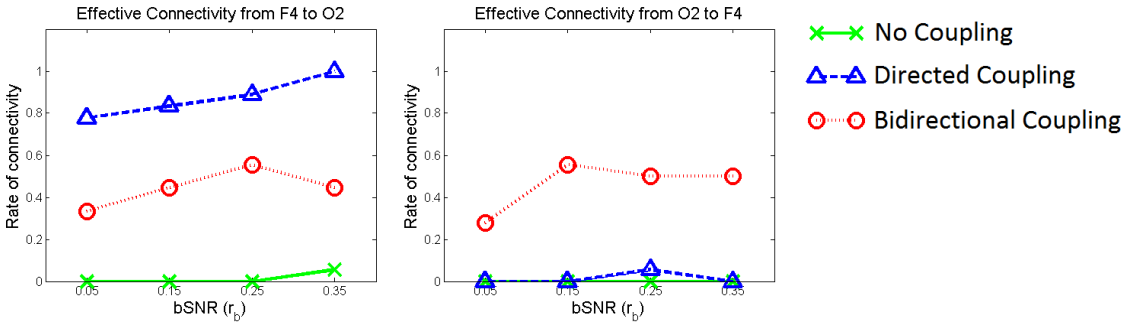

Figure 4.4: Plots of the connectivity rates as a function of bSNR for simulations with no coupling, directed coupling and feedback coupling. In the left panel we see that the PSGC measure has a power to identify a true directed connection ($F4 \rightarrow O2$) which increases with bSNR. The feedback connectivity is identified in roughly 50% of simulations, this decrease in power compared to the directed coupling is a result of using GC analysis. In the case of feedback connectivity, the signals become entangled and the effects cannot be completely removed from the model for the likelihood ratio test. In the both panels, we see very low rates of spurious connectivity in both the no-coupling and directed-coupling cases for bSNR levels of -2.7 dB ($r_b = 0.35$) or lower.

In Figure 4.6 on the left panel, we see slightly less power to identify connections than the subjective threshold in the 'difficult' scenarios (feedback coupling, low bSNR) but increased power to identify directed coupling for the stronger signals. The subjective threshold is lower (less conservative) than the objective threshold, and so it identifies more shifts from noise; this inclusion of additional noise appears to increase the power to identify in marginal cases.

PD Estimator

The top panels, A and B, of Figure 4.7 show the results of the PD estimator with directed coupling ($C_{F4} = 0$, $C_{O2} = 0.015$), and bSNRs of -12.8 db ($r_b = 0.05$) and -4.8 dB

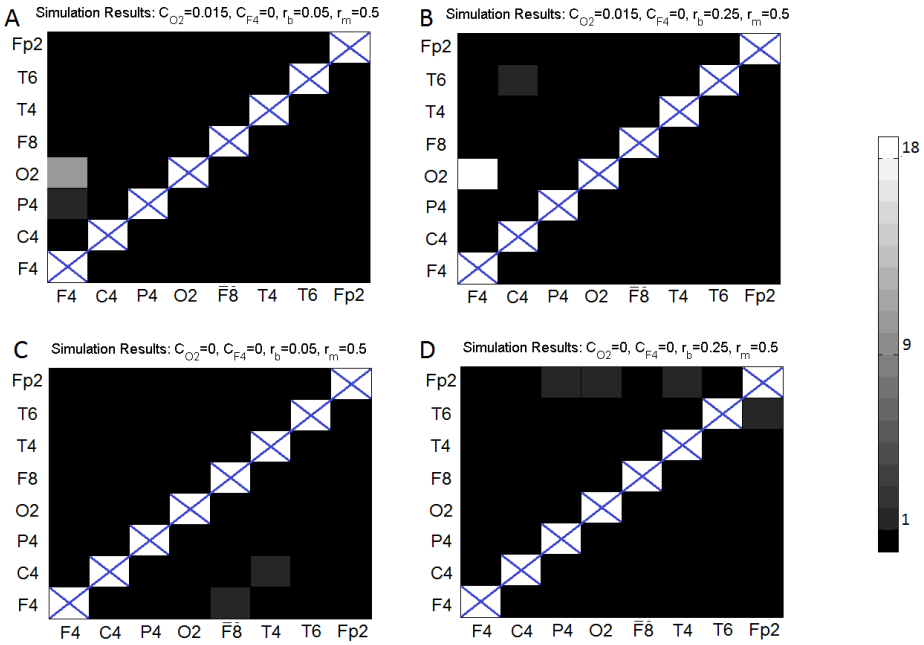

Figure 4.5: Each cell (i,j) of the matrix represents the Granger causality from from N_i to N_j. The shade of the cells (see colourbar) represent the number of simulations which found significant connectivity in the specified pathway. Panels A and B show the results of simulations with asymmetric coupling ($C_{F4} = 0$, $C_{O2} = 0.015$); at bSNRs -4.8 dB (and $r_b = 0.25$) the true connectivity in every dataset ($F4 \rightarrow O2$) and there is only sparse connectivity otherwise. Panels C and D show the results of simulations with no coupling ($C_{F4} = C_{O2} = 0$), where there are minimal spurious connections due to VC.

(and $r_b = 0.25$). The true connection ($F4 \rightarrow O2$) does not stand out from the occasional false positives in the other pairs. At both bSNR levels and both coupling values there are minimal false positive connections; however, given the reduced power to identify the true connection use of this more conservative approach does not appear to reduce the rate of false positives (15, 20%).

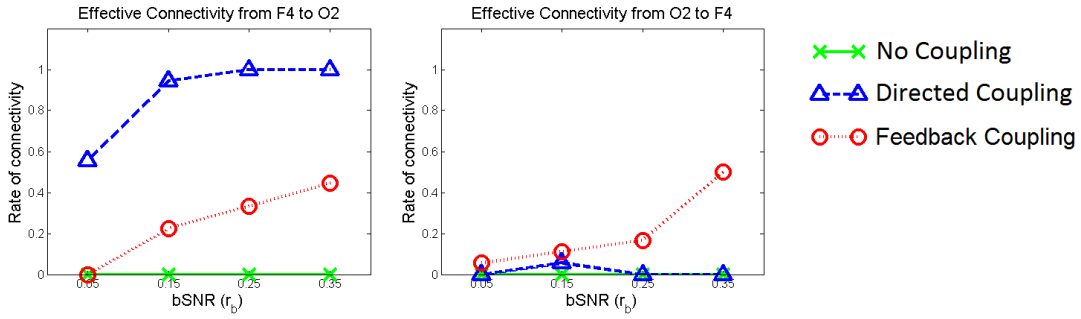

Figure 4.6: Plots of the connectivity rates as a function of bSNR for simulations with no coupling, directed coupling and feedback coupling. In the left panel we see that the PSGC measure has high power to identify a true directed connection ($F4 \rightarrow O2$) increasing to 100% as bSNR increases. The feedback connectivity is also increasing power as bSNR increases, although with lower overall levels due to the difficulty of resolving feedback connectivity with Granger causality. In the both panels, we see very low rates of spurious connectivity in both the no-coupling and directed-coupling cases for bSNR levels of -2.7 dB ($r_b = 0.35$) or lower.

The left panel of Figure 4.8 shows the PSGC measure has power of under 20% to identify the directed connection or feedback for a range of bSNR values. In the right panel, we see the same power to identify feedback connectivity, but only a single spurious connection in both the no-coupling and directed-coupling cases for bSNR levels of -4.8 dB ($r_b = 0.25$) or lower.

4.3.4 Simulation Conclusions

The PD estimator is a conservative in identifying phase shift events, and it was unable to reliably identify connectivity in the simulations. Since the phase shift events undergo additional statistical analyses, a higher rate of false positives may be acceptable if it does not cause errors in the connectivity analysis. For both the objective and subjective threshold selections, there exists a range of bSNR values such that the methods had high power to

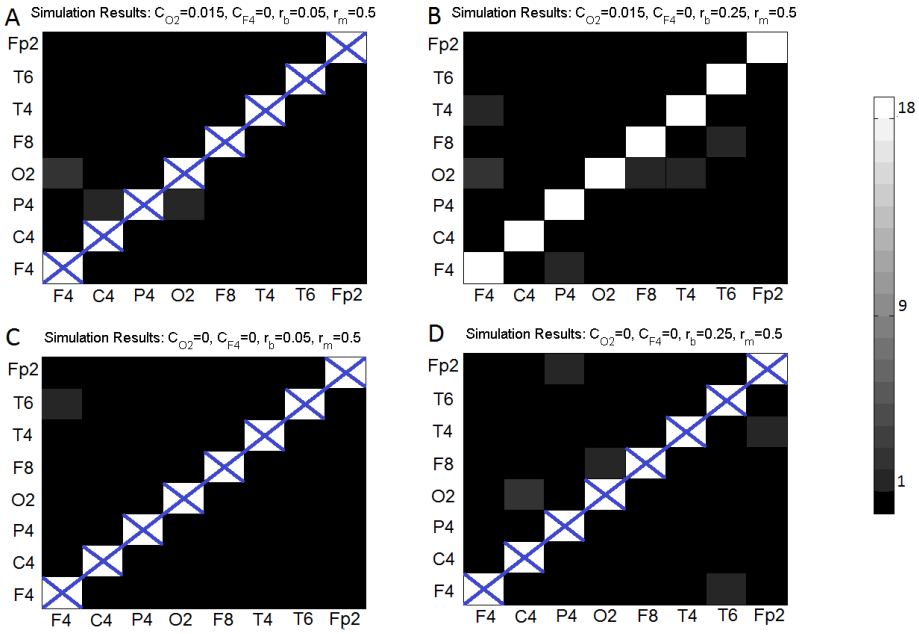

Figure 4.7: Each cell (i,j) of the matrix represents the Granger causality from from N_i to N_j. The shade of the cells (see colourbar) represent the number of simulations which found significant connectivity in the specified pathway. Panels A and B show the results of simulations with asymmetric coupling ($C_{F4} = 0$, $C_{O2} = 0.015$); at bSNRs -4.8 dB (and $r_b = 0.25$) the true connectivity is identified in only 3 of 18 simulations ($F4 \rightarrow O2$) and there is only sparse connectivity otherwise (never in more than 1 of the 18 simulations for a given pathway). Panels C and D show the results of simulations with no coupling ($C_{F4} = C_{O2} = 0$), where there are minimal spurious connections.

identify directed connectivity and were unaffected by VC. A subjective method is undesirable for a number of reasons, and so the objective threshold method is preferred.

For bSNR values greater than -2.7 dB ($r_b = 0.35$), an unrealistically high level for continuous EEG recordings (see Section 4.5), the results show larger amounts of spurious connectivity across many electrodes. This increase in spurious connectivity is likely due to

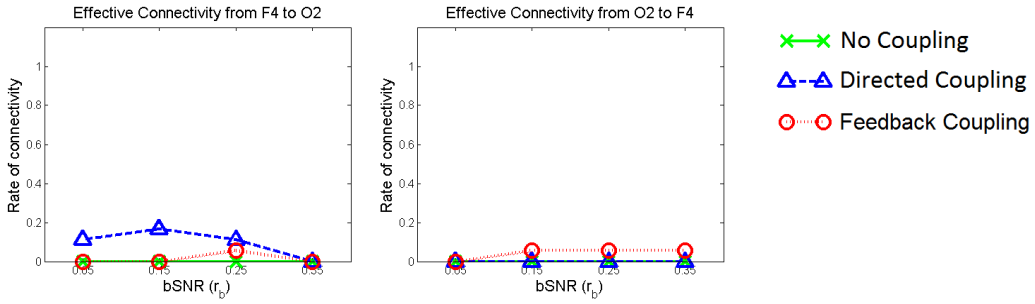

Figure 4.8: Plots of the connectivity rates as a function of bSNR for simulations with no coupling, directed coupling and feedback coupling. In the left panel we see that the PSGC measure with the PD estimator has greatly reduced power, identifying the true connection ($F4 \rightarrow O2$) 20% of the time connection and the feedback connectivity is identified 10% of simulations. This decrease in power compared to the threshold methods is due to the conservative approach of the PD estimator controlling false positives; both true and false positive shift events are identified, both true shifts and false positives.

a combination of VC and average reference effects. The relationship between bSNR and the spread of phase shift events due to VC and average reference is further quantified in Appendix D. As the amount of true signal in the overall activity increases, it starts to dominate the average and shift events in the true signal will be spread across the scalp through the average reference. This result suggests that the PSGC measure is appropriate for estimating connectivity in situations with low bSNR, as is the case with continuous EEG.

In applying the netGC analysis, we found that it successfully identified the true directed connection for each bSNR value $r_b \in (0.05, 0.65)$. There were no significant asymmetries found by the netGC analysis in either the no-coupling or feedback scenarios. In every simulation, the netGC analysis did not show any spurious connectivity due to VC.

If network patterns are to be analyzed, then the LRT (see Equation 2.3) should be used as it can provide statistical validation at the individual level and also identifies both sym-

metric and asymmetric connectivity. For this analysis, we method was able to capture the connectivity and mostly unaffected by VC for bSNR values $r_b \in (0.05, 0.35)$. Importantly, the existence of spatiotemporal noise within the brain (consistent with levels in continuous EEG recordings) helps to attenuate the effect of VC on PSGC in the analysis of individual networks.

4.4 Functional Application

Based on the results of the simulation study, I repeat the application of PSGC to the EEG recordings described in Chapter 2.3 with the objective threshold selection method. For each participant and each ordered pair (i,j), the null hypothesis of no Granger causality from N_i to N_j is tested against the one-sided alternative that there is a Granger causal relationship from N_i to N_j, using a LRT (see §2.2.3). Significance levels were adjusted using the Holm-Bonferroni method [141] to give a family-wise error rate of $\alpha = 0.05$ per hemisphere. Figure 4.9 shows the connectivity patterns for each of the six task (3) by hemisphere (2) combinations. In these figures, each cell of the matrix represents the Granger causality from one neural region to another. The i,j element of the matrix represents PSGC from N_i to N_j. The shade of the cells represent the number of participants who had significant connectivity in the specified pathway. The i,j element is white if causality exists in every participant, black if there is causality in zero participants and shades of grey represent intermediate results. The diagonal entries are crossed out to emphasize that there is no analysis on these locations. There were higher levels of connectivity during the active tasks (auditory, visual) than during the resting condition, this is emphasized in Figure 4.10 where plots of the raw p-values are shown for each task.

Sample connectivity maps from two participants can be seen in Figure 4.11 (connectivity maps for all participants can be found in Appendix F). In the resting condition, there are many short range (neighbouring site) connections, while both the auditory and visual tasks have more long range connections. The connections in the vigilance tasks are pre-

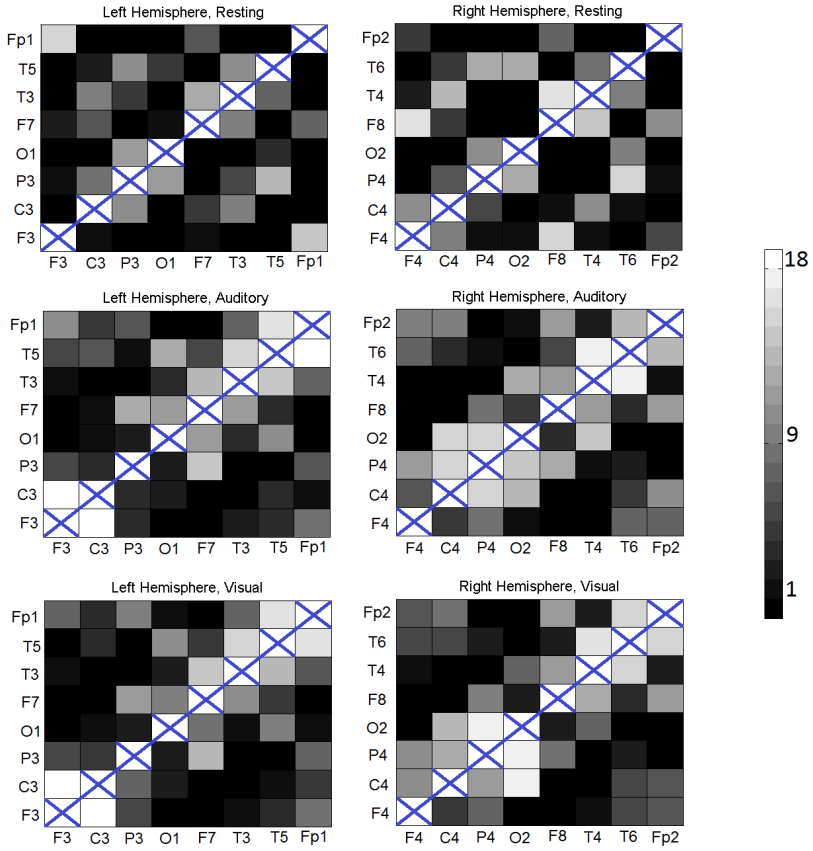

Figure 4.9: Results of the PSGC analysis between electrode sites in 18 EEG recordings. The i,j cell of the matrix represents the PSGC from region i to region j. The grey scale (see colourbar) represents the number of participants where there is significant connectivity. A black cell means no participants had significant connectivity, while a white square signifies information transmission in all 18 participants. Results among left hemisphere locations are on the left, right hemisphere locations on the right. The three rows correspond to resting (top), auditory (middle), and visual (bottom).

dominantly anterior-posterior (front to back), connecting the prefrontal sites (F3/4, F7/8 Fp1/2) to the parietal sites (O1/2, P3/4, T7/8), instead of lateral-medial (side-to-side)

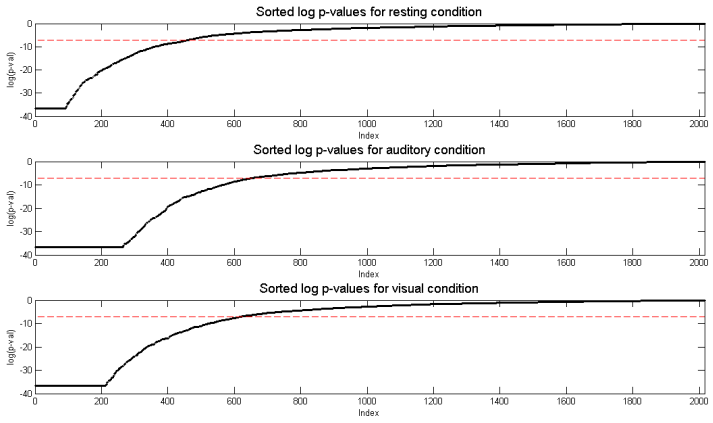

Figure 4.10: Plots of the sorted log p-values for each task condition across all 18 participants. The red line corresponds to a Bonferroni corrected threshold for an $\alpha = 0.05$ significance level. The flat portion at the beginning of the plot are p-values which were essentially zero. From these plots, it can be determined that there are more significant connections than in the activity tasks (auditory, visual) than in the resting task.

connections. This may reflect engagement of the frontal and posterior attention systems in the two vigilance tasks [293].

Taken with the results of the simulation study, the short range connections in the resting task may be a result of localized networks or VC; however, the long range connections in the active tasks would not be observed without some underlying frontal-posterior coupling. The connectivity matrices do not show any obvious group level asymmetries (consistent with the netGC analysis); however, looking at individual networks suggests that asymmetries may exist but are not consistent across participants due to individual differences.

4.4.1 Mean Effects

In addition to the individually estimated networks, we also analyzed task differences in network patterns. The netGC analysis (see §2.2.3) identified only one significant mean asymmetry, a connection leaving occipital sites ($O2 \rightarrow C4$) in the visual condition.

Individual networks were additionally quantified using the network average of local clustering coefficients, a graph theoretical measure which has been used in describing 'small-world' networks in functional connectivity [263, 295]. The local clustering coefficient measures the degree to which neighbours of a node are interconnected. The average clustering coefficients were used as the dependent variable in a 2 (hemispheres) x 3 (conditions) repeated measures ANOVA, with p-value corrections for sphericity [122]. The resting task ($\hat{\mu} = 0.114$, $\hat{\sigma} = 0.028$) had significantly lower clustering coefficients than the auditory ($\hat{\mu} = 0.243$, $\hat{\sigma} = 0.047$) and visual tasks ($\hat{\mu} = 0.196$, $\hat{\sigma} = 0.031$), $F(2, 34) = 5.213$, $p = 0.022$.

A further test of the PSGC effects by examining the directed connection patterns as a function of the three conditions across participants. The number of significant directed connections per person was the dependent variable in a 2 (hemispheres) x 3 (conditions) x 2 (connection lengths) within-subject analysis of variance, with p-value corrections for sphericity using Greenhouse-Geisser. The right hemisphere sites ($\hat{\mu} = 8.417$, $\hat{\sigma} = .381$) demonstrated significantly more overall connections than the left ($\hat{\mu} = 7.176$, $\hat{\sigma} = .384$), $F(1, 17) = 12.374$, $p = .003$. Similarly, the vigilance tasks elicited more connections ($\hat{\mu} = 8.958$, $\hat{\sigma} = .684$ and $M = 8.347$, $\hat{\sigma} = .381$ for the auditory and visual tasks, respectively) than did the resting condition ($\hat{\mu} = 6.083$, $\hat{\sigma} = .349$), $F(2, 34) = 14.308$, $p < .001$. There were significantly more short ($\hat{\mu} = 9.139$, $\hat{\sigma} = .431$) than long ($\hat{\mu} = 6.4537$, $\hat{\sigma} = .394$) connections, $F(1, 17) = 55.63$, $p < .001$. There was a Hemisphere x Connection Length interaction, $F(1, 17) = 8.011$, $p = 0.012$; both hemispheres showed similar amounts of long connections, but the right hemisphere had more short range connections ($\hat{\mu} = 8.037$, $\hat{\sigma} = .403$) than the left hemisphere ($\hat{\mu} = 10.241$, $\hat{\sigma} = .551$). Of particular interest was the Task x Connection Length interaction, which was highly reliable, $F(2, 34) = 145.5$,

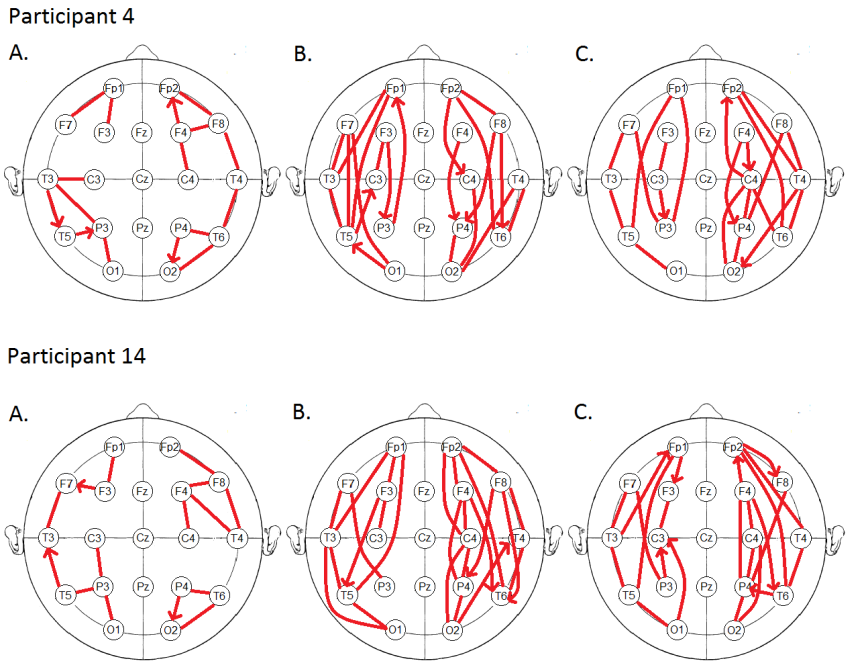

Figure 4.11: Maps of the estimated network connectivity during the A: resting, B: auditory, and C: visual tasks, for the LRTs of two individual participants. In the resting task we see many short range connections but relatively few long range connections. In contrast, the active tasks show many connections over longer distances, primarily anterior-posterior, connecting prefrontal sites to parietal sites. Asymmetries are observed in these individual networks; however, the direction is not consistent between individuals. This pattern was consistent across all participants and is further quantified in Figure 4.12.

$p < .0001$, with the mean values illustrated in Figure 4.12. To examine the reliability of this pattern further, we constructed for each participant a metric reflecting the extent to which the resting condition was associated with more short connections than the active tasks and the active tasks more long connections than the resting condition. This yielded totals for the left hemisphere and right hemisphere for each participant. Each hemisphere of every participant followed this general pattern ($\hat{\mu} = 9.5$, range = 0.5 to 19.5, with any value above zero indicating this general pattern), indicating extremely high reliability of

the effect illustrated in Figure 4.12.

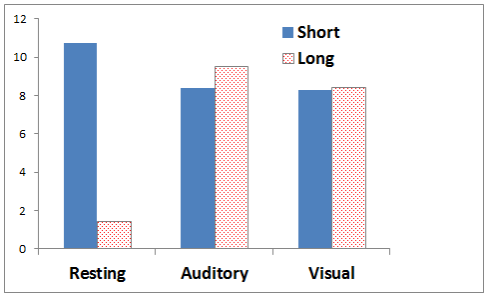

Figure 4.12: The significant Task x Connection Length interaction, demonstrating an increase in longer connections and a reduction in shorter connections during the auditory and visual vigilance tasks compared to the resting condition.

4.5 Discussion

This chapter demonstrates a novel method for analyzing neural connectivity in a simulation study mimicking the properties of EEG recordings. From the results of the simulation study, the netGC analysis is the preferred test for mean asymmetric connectivity; it is designed to increase the statistical power to identify asymmetries and the results of the simulation study suggest that it is not affected by VC for a large range of bSNR values.

Calculating the bSNR from EEG recordings requires the use of an inverse method to identify the different biological components, making estimates of the bSNR highly dependent on the specific inverse method used. To get an idea of the bSNR that can be expected in non-averaged EEG data, we can use results from independent component analysis (ICA). In [72], twenty different ICA algorithms are each applied to EEG recordings from a visual attention task; they found an average of 22.4 dipole-like biological components (range 9.1 - 37.4). The strengths of the individual components are not reported, but we can infer an average component strength of 4.5% (range 2.6-11%). This is well below the $35/2 = 17.5\%$

per component level where VC and reference electrode became a concern in the simulations. A bSNR of greater than -2.7 dB is not realistic for non-averaged EEG recordings which are notorious for having large amounts of noise, so we conclude that observed connectivity difference between resting and vigilance tasks are not due to VC or reference electrode.

In the simulation study, it was confirmed that the PSGC measure is able to correctly identify the direction of information transfer in conditions similar to those of EEG recordings. The simulated data contain multiple sources of noise, both independent additive noise and spatiotemporal sources of biological noise. Results of the simulation showed that the existence of spatiotemporal noise helped to attenuate the spread of phase shift events due to VC or average reference. The LRT on individual networks did not show spurious connectivity when the bSNR levels were below -2.7 dB but larger bSNRs showed spread of connectivity across several electrodes. The ability to identify connectivity in scenarios with very low bSNR is a strength of the PSGC measure; the high levels of noise that are typical in continuous EEG recordings actually help reduce the extent to which VC produces spurious connectivity measurements. Intuitively, this is because in a mixture of signals, the phase of the mixture is a weighted sum of the phase of individual components. This means that only the strongest components of the mixture will be able to induce a shift event. Since the affect of volume conduction decreases with distance, the more biological noise that exists in a system the less likely a generator is to be the strongest component of the EEG signal in locations away from the source.

The existence of high levels of noise is one of the main challenges in EEG analysis. The most popular method of dealing with this noise is to average time-locked waveforms from many trials [200], under the assumption that unrelated background activity will decrease as the number of trials goes up; however, evidence suggests that phase information contained in the background activity is important to understanding brain activity [181]. This emphasizes the need for measures like PSGC which can be used in low bSNR conditions to analyze non-averaged EEG recordings.

When analyzing a mixture of signals, the overall phase is an amplitude weighted func-

tion of the phase of the individual components. This means that the phase of the dominant component is most represented in the signal. When identifying phase shift events, contributions from weaker components of the signal may not have a large enough effect on the overall phase to reach the threshold value. The use of a well chosen threshold value can effectively eliminate phase shifts from weak contributions, such as those due to VC over large distances (strength of VC decreases quadratically with distance). This result is confirmed by the simulation study, where low bSNRs (≤ -2.7 dB) showed only sparse effects of VC, while the high bSNR (> -2.7 dB) showed high rates of observed connectivity due to VC. Further evidence for this result is shown in Appendix D, investigating the effects of noise, VC and reference electrode on the identification of phase shift events.

Another consideration for use in EEG recordings is the effect of reference choice on the phase of a signal. Using an average reference can cause inflated connectivity in measures such as coherence [205, 229], and can cause phase variables to lose physiological importance [275]. The latter is not the case when measuring PSGC, because it analyzes qualitative features of the signal (phase shifts modelled as a point process) rather than interpreting the magnitude of the phase variables directly. Additionally, because each signal contributes only a small amount to the average (1/121 in the real recordings), with realistic amounts of biological activity the phase shifts in the reference will be strongly attenuated and therefore will be eliminated by the threshold technique.

4.5.1 Conclusions and Future Directions

The PSGC measure is a novel method for quantifying the connectivity between electrode sites, a method which is strengthened by its ability to detect directed (or asymmetric) connectivity, its high temporal resolution, and its ability to detect non-linear relationships between electrodes. A simulation study was used to assess the effect of VC on the PSGC measure, and the results suggest that the identified connectivity patterns in the EEG application cannot be attributed to VC or the use of an average reference.

The results of this simulation study have been packaged together with the material from Chapter 2 to provide a proof-of-concept demonstration that the PSGC measure is appropriate to use in conditions similar to those of EEG recordings.

The simulation study suggests that the PSGC measure is resilient to VC in appropriate circumstances; however, there are situations where VC may have an effect. It was assumed for simulation that neural sources are perpendicularly oriented, and while this is a reasonable assumption (see §1.3), tangentially oriented sources do occur [299] and may confound PSGC results. It is also not possible to precisely localize the source of neural connectivity from the scalp recordings. A natural extension of the current PSGC method which may overcome these limitations is to modify it for data derived from independent component analysis (or source modelling techniques). An investigation into such an extension is explored in Chapter 5.

Chapter 5

Reconstruction of Phase Shift Events

In Chapter 4, a simulation study was used to show that the interpretation of phase shift Granger causality (PSGC) analysis as neural connectivity, relies critically on the existence of substantial spatio-temporal 'biological' noise which obstructs the spread of phase shift events. Although evidence is shown that electroencephalogram (EEG) recordings typically have the required noise, it is not possible to ascertain if it is true for any specific recording.

In this Chapter, I investigate the possibility of applying PSGC directly to reconstructed neural sources rather than scalp recordings. If this can be accomplished successfully, it would greatly increase the spatial resolution of the PSGC method, and provide additional certainty that observed connectivity was not a result of volume conduction (VC). The first step towards such an analysis is to determine which reconstruction techniques (if any) are best suited for localizing the neural source of EEG phase shift events.

In this chapter, my contribution to this field of research is a detailed simulation study used to determine the extent to which different source reconstruction (SR) techniques are able to provide a spatio-temporal estimate of phase shift events. In addition, the simulation study is used to assess the ability of SR techniques to localize neural sources, results typically depend on the particular situation, and so there is no established consensus in the literature on the optimal technique.

5.1 Introduction

The effect of VC on connectivity metrics can be reduced by estimating the connectivity between neural sources rather than scalp recordings, though it can never be completely removed [249]. There are two aspects to estimating neural connectivity from scalp recordings; localizing the sources of neural activity and then estimating the connectivity between sources.

Traditionally, localization is performed using SR techniques (see §1.6) and then connectivity measures (see §1.5) are applied to time series of reconstructed source activity. The most popular choice is to use the minimum norm estimate (MNE) for source reconstruction [8, 11, 14, 63, 132, 149], while there has been more variation in the choice of connectivity measure, including coherence [132, 149], Granger causality measures such as the directed transfer function (DTF) and partial directed coherence (PDC) [8, 11, 14], or phase synchronization [46, 149]. The optimal combination depends on the relationship between sources and whether the coupling is linear/non-linear, narrow-band/broadband, and the more general the metric the more data is required for good estimates and more tuning parameters to be selected [63]. A specific comparison between DTF and PDC measures of Granger causality finds that the PDC measure has less bias but greater variance than the DTF measure when estimating source connectivity [10].

When estimating source connectivity using distributed SR techniques (i.e. MNE) with many voxels, a crucial aspect is the selection of *regions of interest* (ROI) for connectivity analysis. The selection of ROIs is non-trivial, and many authors elect to use subjective *a priori* ROIs based on the physiology of the problem [8, 11, 14, 132]. One technique is to perform preliminary coherence analysis with a peripheral biological signal (such as heart/muscle activity) [149], but an appropriate signal is not always available and this method may be misleading if an arbitrary signal is used [249]. Time coherent expansion is an algorithm which uses iterative expansion and contraction of neural regions to define ROIs with temporally coherent signals [64], this method still requires an initial seeding of expected regions and does not completely solve the ROI problem. A promising approach

is to use combined EEG/fMRI (functional magnetic resonance imaging) brain imaging to make use of the superior spatial resolution of fMRI signals and define ROIs [14], although this is not always feasible due to the increased technical requirements for fMRI or joint EEG/fMRI recordings.

Modern statistical methods such as independent component analysis (ICA) allow source activity to be estimated before (or concurrently with) source localization. It is somewhat counter-intuitive to estimate source connectivity using ICA, as the ICA algorithm is attempting to identify maximally independent sources, despite this, there have still been positive results from this approach [180, 181, 207, 73]. A modification of this approach is to perform joint multivariate auto-regressive (MVAR) estimation and ICA demixing to separate systematic coupling between sources and independent residual terms [77, 112]. These algorithms are improved in [137] by including an appropriate least absolute shrinkage and selection operator (LASSO) sparsity penalty on a subset of the parameters.

Dynamic Causal Modelling (DCM) is a Bayesian modelling framework for estimating effective connectivity between neural regions, originally introduced in the fMRI literature [98]. Extensions of this idea are employed in EEG analysis [157], which use a hierarchical Bayesian model of the electrical activity produced in the brain. Like the ICA algorithms, the DCM method provides joint estimation of source reconstruction and source connectivity parameters in a single step.

In both standard Bayesian SR and Bayesian DCM, a variational Bayes (VB) approximation to the posterior distribution is often used to estimate parameters [61, 97, 158, 157, 185, 203, 219, 243, 282]. This approximation greatly speeds up the calculation of a posterior distribution, especially when compared to Markov Chain Monte Carlo (MCMC) sampling solutions. The VB approximation is able to provide accurate point estimates of the parameters; however, in Appendix C we show it to greatly underestimate variability in situations with high covariance between parameters in the true posterior. Despite this limitation, the VB method is still actively used in Bayesian source reconstruction.

In this chapter I perform a simulation based studies to compare the ability of source connectivity techniques to accurately localize phase shift events. Phase shift events were generated from coupled Rössler attractors, and then linearly mixed to EEG sensors using realistic lead field matrix. Several methods of source connectivity analysis were implemented to determine which is most accurate at phase shift localization.

5.2 Methods

Several SR techniques are used to localize phase shift events. Due to the underdetermined nature of inverse problem, regularization techniques are used to determine a unique solution. I consider three different penalty functions for regularization of distributed solutions to the inverse problem, as well as ICA and Bayesian statistical approaches.

5.2.1 Distributed Solutions

The distributed approach for solving the EEG inverse problem is to consider the generation of S scalp potentials from a set of sources on a pre-specified lattice of G volume elements (or voxels). Distributed methods typically include much greater number of voxels than sensors ($G >> S$) which make-up the solution space, resulting in an a highly underdetermined inverse problem. For this application, a solution space is created which constrains the voxels to the surface of the brain; not all electrical brain activity originates at the surface, but due to its proximity to the scalp, the perpendicular orientation of pyramidal neurons, and the synchronization of clusters of neurons, it is reasonable to expect that surface activity dominates the EEG recording in many applications (see §1.3).

To create the solution space, first a regular 3D lattice (with inter-voxel distance of 0.01) is made to fill the entire head, and then a spherical shell is created by keeping only voxels with $z > z_0$ (to remove the 'bottom' of the brain) and a distance between $r = 0.869$ and $r = 0.872$ units from the origin (representing the surface of the brain, see §5.3.1). This results in G=292 voxels, shown in Figure 5.1 along with the locations of the generators

used in the simulation study.

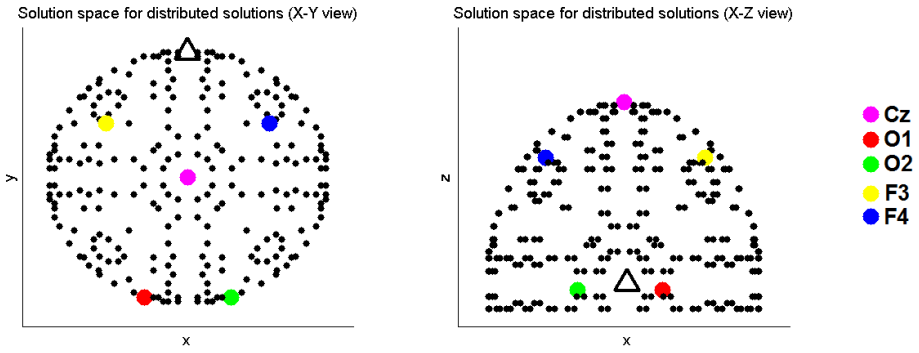

Figure 5.1: Solution space for distributed source reconstruction techniques. Each blue point indicates of voxel of the solution space, the nose is indicated by the black triangle. From the top-down view (X-Y), the coupled oscillators are on opposite corners, and the uncoupled source is in the centre of the head.

The lead field matrix (K) is calculated using the a 3-shell approximation [23] (see Section 1.6.1). When solving the source reconstruction problem, it is crucial to consider the reference electrode which has been used. Here we have used the average reference, so we update the lead field matrix to account for this average reference (K_{AR}),

$$K_{AR} = (I_S - 1_S 1_S'/S)K,$$

where I_S is the S by S identity matrix, and 1_S is a S by 1 vector of ones.

Minimum Norm Estimate

The MNE is one of the earliest methods of source reconstruction [144, 131], partially due to its computationally feasibility and the existence of a tractable solution. For observed scalp potentials Y, the MNE solution seeks the source configuration J which minimizes both the least squares penalty and the ℓ_2 norm of the source estimate,

$$\hat{J} = \min_J ||Y - K_{AR}J||_2 + \alpha_{MNE}||J||_2.$$

where α_{MNE} is a tuning parameter determine the smoothness of the estimated current. The solution to this problem [214] is

$$\hat{J}_{MNE} = K'_{AR}(K_{AR}K'_{AR} + \alpha_{MNE}H)^{+}Y,$$

where X^{+} denotes the Moore-Penrose pseudo-inverse of X.

sLORETA

The MNE has been shown to be biased towards superficial sources [214]. A correction to overcome the depth bias of the MNE, the standardized LOw Resolution TomograAphy (sLORETA) algorithm 'standardizes' the sources by an estimate of spatial variance. The MNE estimate can be shown to be equivalent to a Bayesian formation of the inverse problem

$$Y = K_{AR}J + E,$$

where J are the current density parameters, E is independent Gaussian noise with constant variance

$$\Sigma_E = \alpha_{MNE}I_{3G}$$

and the prior distribution on J is also Gaussian with mean zero and variance

$$\Sigma_J = I_{3G}.$$

Under this model, the variance of the source estimator is given by [215]

$$S_{\hat{J}} = K'_{AR}(K'_{AR}K_{AR} + \alpha_{MNE}H)K_{AR}$$

The solution at voxel g is then [215]

$$\hat{J}_{g,sLORETA} = \hat{J}_{g,MNE}\left([S_{\hat{J}}]_{gg}\right)^{-1}\hat{J}'_{g,MNE}$$

where $\hat{J}_g$ is the 3×1 current density vector at voxel g, and $[S_{\hat{J}}]_{gg}$ is the g^{th} 3×3 block diagonal matrix of the estimated spatial variance $S_{\hat{J}}$.

LASSO

A sparse estimate of the current density can be achieved by replacing the ℓ_2 penalty in the MNE with the ℓ_1 (or LASSO) penalty [278],

$$\hat{J_{LASSO}} = ||Y - KJ||_2 + \alpha_{LASSO}||J||_1,$$

where α_{LASSO} is a tuning parameter determining the sparsity of the estimated currents. A sparse solution will emphasizes focal sources of neural activity, as opposed to the smoother solutions given by the MNE. The Least Angel Regressor Selection (LARS) algorithm is used to solve for $\hat{J}_{LASSO}$ [79].

5.2.2 Bayesian Hierarchical Model

With the Bayesian approach to source reconstruction, a 2-level hierarchical regression model is used similar to ones used in [282]. The response variable $Y_{S\times 1}$ is a vector of electric potentials recorded from S sensors. The top level of the model relates Y to the parameters of interest $J_{3G\times 1}$, a matrix of primary current densities from G generators,

$$Y|J = K_{AR}J + \epsilon_Y.$$

The S by 3G matrix K_{AR} is the known LFM which is calculated by solving the EEG forward problem, the term ϵ_Y represents independent measurement errors at the sensors.

In the second level of the model describes the current density vector J,

$$J = \epsilon_J.$$

The currents are modelled as having zero mean and the term ϵ_J represent the source innovations. The errors ϵ_j are also independent from ϵ_Y.

Prior Distribution

The error terms in the hierarchical model are assumed to follow a multivariate Gaussian distribution with mean zero. For the top level of the model, the errors are independent

across sensors, a vector $\sigma_{S\times 1}$ determines the precision for each sensor

$$\epsilon_Y \sim G\left(0, diag(\sigma)^{-1}\right).$$

The errors in the second level are independent across generator, the vector $\lambda_{3G\times 1}$ allows the precision to be different at each generator and in each direction,

$$\epsilon_J \sim G\left(0, diag(\lambda)^{-1}\right).$$

The prior distribution for the precision hyper parameters will use independent gamma distribution, each with their own hyper parameters,

$$p(\sigma) = \prod_{i=1}^{S} Ga(a_{\sigma,i}, b_{\sigma,i}),$$

$$p(\sigma) = \prod_{i=1}^{3G} Ga(a_{\lambda,i}, b_{\lambda,i}).$$

This choice of prior results in conjugate normal-gamma prior distributions for mean-precision parameters. The use of gamma priors for the precision variables is also called *automatic relevance detection* (ARD) priors [198], ARD priors allow the data to determine which generators and sensors (corresponding to λ, σ) are most relevant to the model. Following [282], the values of the hyper parameters in the ARD priors were taken to be $a = \frac{1}{1000}$ and $b = 1000$. This results in an uninformative prior distribution which has expected value of 1 and a variance of 1000, allowing the data to dominate the estimates of precision parameters.

Variational Bayes Approximation

Variational Bayes is a parametric method for approximating the posterior distribution in Bayesian analysis. The method is typically applied in situations where the complex nature of the problem makes analytic solutions difficult or impossible and sampling methods such as MCMC are too slow. The idea of VB is to restrict attention to candidate distributions which can be factored into groups of parameters and from these candidates

selecting the approximate posterior which maximizes the *variational free energy*. With an appropriate factorization of parameters, this often leads to an analytically tractable and efficiently computed solution. Optimization is done using an process similar to the expectation-maximization algorithm, described in Appendix C.

For the current model, the posterior for σ is a Gamma distribution and the update equations for the VB approximation are

$$\begin{aligned} a_{\sigma,i} &= a + 0.5, \\ b_{\sigma,i} &= \left(\frac{1}{b} + \frac{1}{2}(Y_i - K_{\cdot,i}J)^2\right)^{-1}, \\ \sigma_i &= a_{\sigma,i} b_{\sigma,i}. \end{aligned}$$

Similarly, the precision parameters for the generators, λ, also has a Gamma posterior, and update equations

$$\begin{aligned} a_{\lambda,i} &= a + 0.5, \\ b_{\lambda,i} &= \left(\frac{1}{b} + \frac{1}{2}J_i^2\right)^{-1}, \\ \lambda_i &= a_{\lambda,i} b_{\lambda,i}. \end{aligned}$$

The current density parameters have a Gaussian posterior, and the update parameters are

$$\begin{aligned} \Sigma_J^{-1} &= diag(\lambda) + K' diag(\sigma) K, \\ \mu_J &= (Y' diag(\sigma) K)\Sigma_J. \end{aligned}$$

5.2.3 Independent Component Analysis

An ICA algorithm attempts to find the linear transformation which results in a set of maximally independent signals. There are many different ways to define a cost function

which quantifies the independence of a set of signals, and thus there are many different implementations of the ICA method. In [72], twenty different ICA algorithms are compared in their ability to localize neural sources, and the most successful algorithms were based on the infomax principal [171]. The infomax principal suggests to find the linear combination which maximizes the entropy of J (see §1.5.3). For the current application, the *runICA* algorithm in EEGLAB [71] is used which is based on the Infomax principal. Once components have been identified, scalp projections of components are compared to potential dipoles in the solution space, and components which account for sufficient variance are fit as dipoles while others are rejected as noise components. This is process is performed using the dipFIT plug-in in EEGLAB [208].

5.3 Experiment

The simulated data for this experiment come from systems of coupled Rössler attractors, as described in §3.4. There are a total of 5 attractors, two pairs of weakly coupled oscillators displaying phase shift events and a fifth uncoupled oscillator. The first pair of weakly coupled oscillators are generated with $\omega_{1/2} = 2\pi f_0 \pm \delta\omega$, $f_0 = 9$, $\delta\omega = 0.0675$ and $C = 0.12$. The second pair of oscillators is generated with similar values except that $f_0 = 5$ and $C = 0.06$, and the uncoupled oscillator is generated with $f_0 = 2$ and $C = 0$.

Generators are located on the surface of the brain (shown in Figure 5.1); the first pair of coupled oscillators are located under scalp regions F4 and O1, the second pair under F3 and O2, and the uncoupled oscillator under Cz. Each generator is oriented perpendicular to the surface of the brain, and the amplitude is taken to be the x_t time series of the corresponding Rössler attractor, a 20.83 minute signal sampled at a rate of 8Hz, for 10000 data points. Figure 5.2 shows the time series of amplitudes for each generator, along with the corresponding estimate of their PSD. Both the F4 and O1 generators show a distinct frequency peak at 1.3 Hz, while the F3 and O2 generators show a peak at 2.8 Hz. The estimated phase difference for each pair of coupled oscillators in the (1-1.5) Hz and (2.5-3) Hz frequency bands is shown in Figure 5.3. From the plots, it is seen that there are

four phase shift events between the F4 and O1 oscillators are at {2.77, 6.06, 11.77, 15.98} minutes and six phase shift events between the F3 and O2 oscillators at {1.81, 5.04, 7.96, 9.63, 11.71, 15.04} minutes.

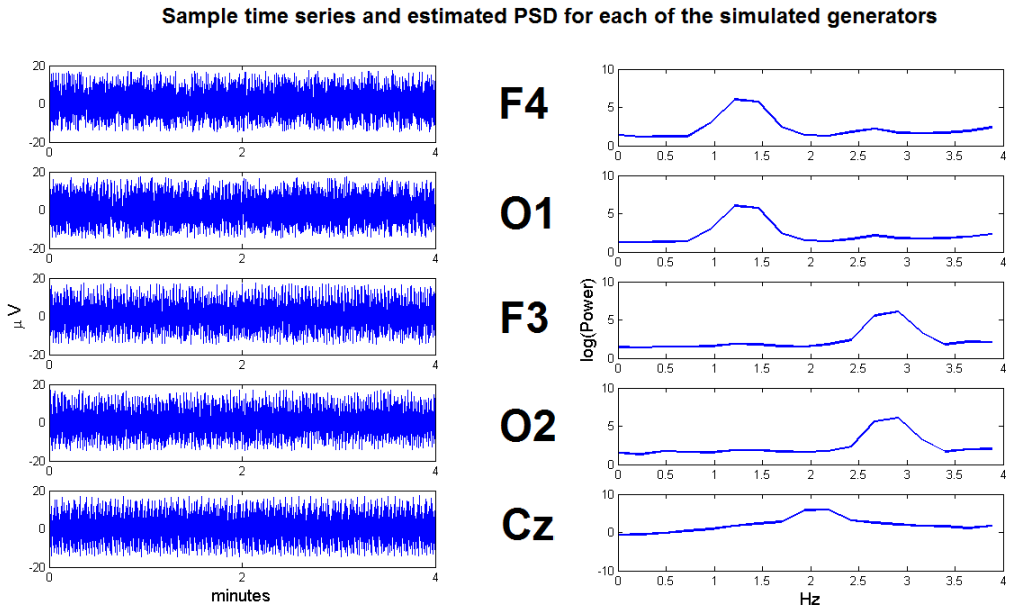

Figure 5.2: Sample time series and estimated power spectral densities for each of the simulated generators. The are no visible changes in the amplitudes of the oscillators. The top two plots (F4 and O1) show a distinct frequency peak at 1.3 Hz, the third and fourth plots (F3 and O2) have a peak at 2.8 Hz and the bottom plot of the uncoupled oscillator has a peak at 2Hz.

5.3.1 Head Model

To emulate the field spread due to VC associated with EEG recordings, the generated sources ($J_{3G\times 1}$) are linearly mixed to $S = 19$ scalp electrodes ($Y_{S\times 1}$) in the 10-20 system

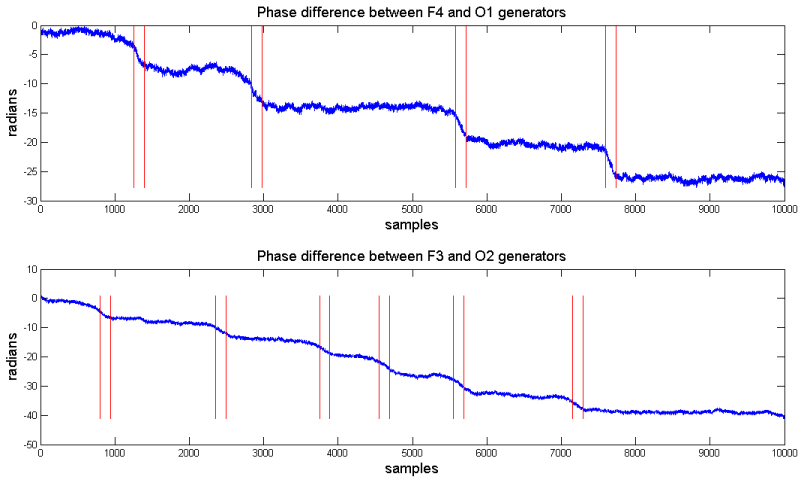

Figure 5.3: True phase difference of coupled Rössler attractors at F4-O1 (top) and F3-O2 (bottom). Both systems show distinct periods of synchronized phase activity and spontaneous desynchronizations. The phase shift events are more obvious in the top panel than the bottom, but can be clearly identified in both situations. There are four phase shift events in the F4-O1 system and six phase shift events in the F3-O2 system.

using a mixing matrix ($K_{s\times 3G}$),

$$Y = KJ.$$

The LFM, K is computed from the geometric and conductive properties of the head model. In this work, the three concentric sphere head model is used, with the surface of each sphere representing the brain, skull and scalp. Following [239], the radii of the brain, skull and scalp are 0.87, 92, and 1.00 respectively with corresponding conductivities of 0.33, 0.0041, 0.33. To calculate the LFM, the 3-shell model is approximated as the sum of three appropriately scaled 1-shell models [23] (see Appendix A).

The signal-to-noise ratio (SNR)s is set by taking a weighted average of the noise-free data (Y) and independent measurement noise (ϵ) representing measurement error at the

sensor,

$$Y^* = \frac{rY}{||Y||} + \frac{(1-r)\epsilon}{||\epsilon||}.$$

The relationship between r and the SNR is given by $SNR = 10\log_{10}\frac{r^2}{(1-r)^2}$ decibels.

The generated sources are mixed to 19 scalp locations of the 10-20 system of electrode placement (see Figure 1.7). Scalp data were initially referenced to site Cz, but then each channel was post-processed, after noise was added, to be re-referenced against the scalp average,

$$Y^*_{AR} = Y^* - \overline{Y^*}.$$

For the electrodes which are directly above the generators, Figure 5.4 shows the final signals after post-processing (Y^*_{AR}), along with their estimated power spectral densities (PSDs). Both pairs of scalp signals corresponding to coupled generators have a frequency peak in the correct location. There are several minor secondary peaks corresponding to the frequency of distant generators, this is an example of the effect of volume conduction. The scalp signal above Cz does not have a frequency peak in the expected location (2Hz), but rather only two small peaks corresponding to the pairs of coupled oscillators. This suggests that the spatial and temporal properties of the generators are having a canceling effect on the activity from Cz; this is in agreement with result in Figure 4.1, that additional spatio-temporal noise attenuates the spread of activity. Scalp topographies corresponding to the three main frequency components in the simulated data; the topographies of the 1.3Hz and 2.8Hz components show the signals as being strongest in the correct locations, while it is not possible to correctly identify the location of the 2Hz generator from the topography.

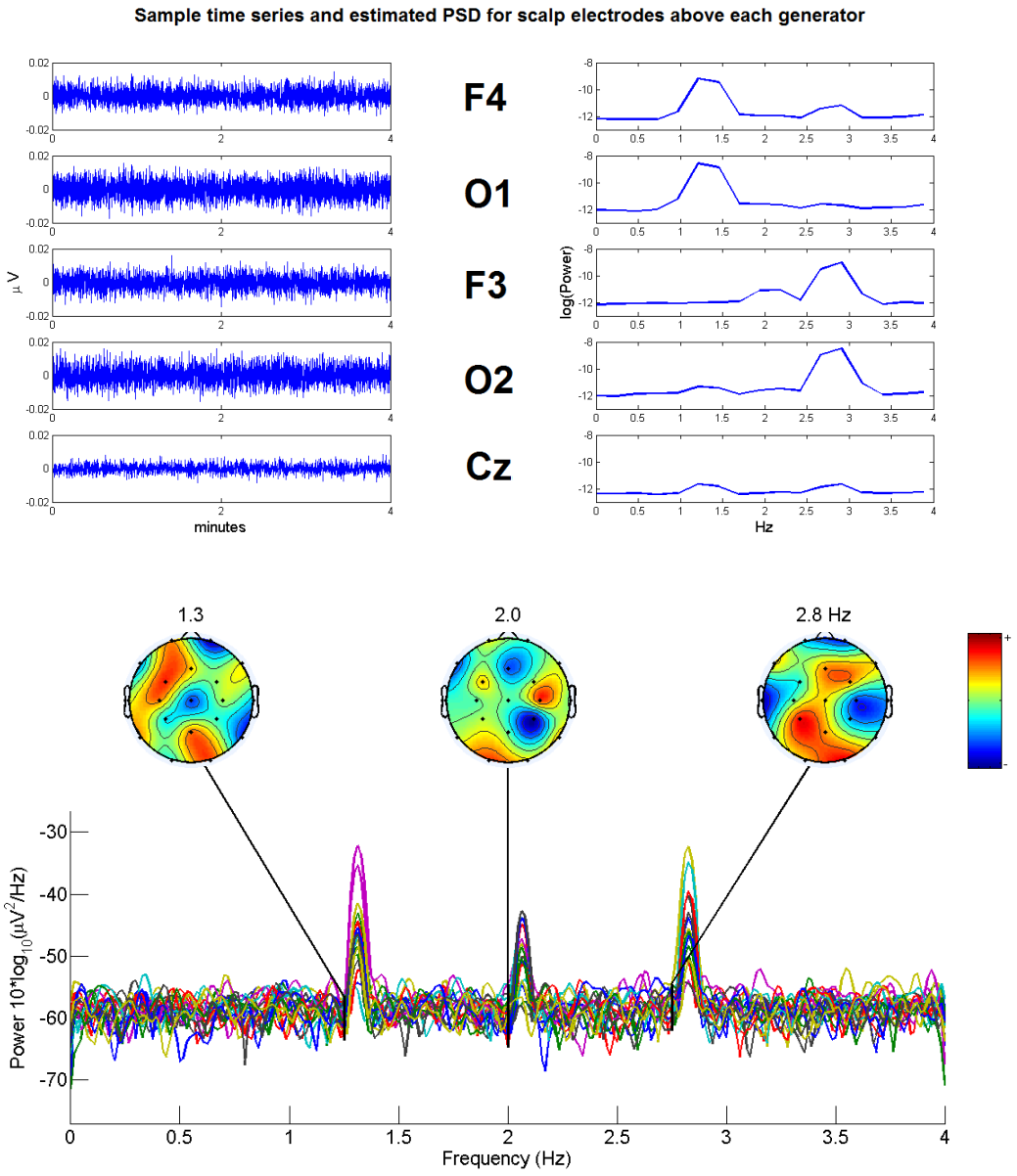

Figure 5.4: **Top**: Simulated scalp signals along with corresponding PSD estimates. Both pairs of scalp signals corresponding to coupled generators have a frequency peak in the correct location, and there are several minor secondary peaks corresponding to the frequency of distant generators (an effect of VC). The signal at Cz does not have a frequency peak in the expected location (2Hz), but rather only two small peaks corresponding to the pairs of coupled oscillators. **Bottom**: Scalp topographies corresponding to the three main frequency components in the simulated data; it is possible to localize the couple generators from the topographies of the 1.3Hz and 2.8Hz components, while it is not possible to correctly identify the location of the 2Hz generator at Cz.

5.3.2 Application Details

For this application, ROI selection is performed using an *a priori* selection of the number of sources as well as a preliminary seeding of source locations. The source space is partitioned into regions, with each voxels going into the region corresponding to the closest source seed. For each region, a source is estimated at the voxel with the largest magnitude.

Solutions to the source reconstruction problem are calculated for a range of tuning parameter values, and the parameter value is selected which minimizes the mean squared localization error from all 5 estimated sources. This process results in parameters values of $\alpha_{MNE} = \alpha_{sLORETA} = 10^{3.5}$ and $\alpha_{LASSO} = 10^{-3}$ (note that the LASSO problem uses the dual formation of the optimization, so the interpretation of these two different values is actually similar).

The parameter τ which defines the auto-correlation in the estimated phase difference signals, must be estimated before the shift identification algorithms can be applied. For each pair of signals subject to the analysis, the phase difference is calculated without applying the unwrapping algorithm; this reduces any bias in the auto-correlation function (ACF) due to phase shift events. The value of τ is taken to be the first zero crossing of the ACF for the corresponding signals. Typical values of τ for this analysis ranged from 16-30 time points.

5.3.3 Localization Results

Each of the distributed methods is able to correctly localize the two pairs of coupled oscillators, but all have trouble localizing the uncoupled oscillator at Cz. Figure 5.5 shows the solution space with different symbols where the source location is estimated for each method. The ICA algorithm provides the best estimate of the source at Cz, while the sLORETA algorithm performs the best among the distributed solutions. This is emphasized in the mean localization errors shown in Table 3.1. The VB and LASSO methods

give the worst estimates of the Cz generator.

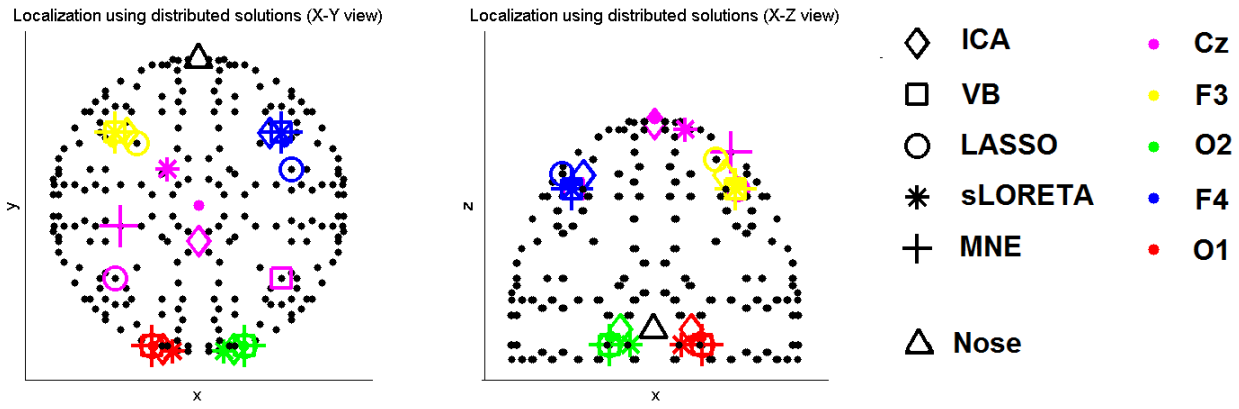

Figure 5.5: A top-down (X-Y, left) and head on (X-Z view, right) view for the results of source localization. Source reconstruction techniques are differentiated by various symbols, specific generators are denoted using colours. The black triangle is at the front of the head model, and represents a 'nose'. Once properly tuned, each of the source reconstruction techniques is able to correctly localize the two pairs of coupled oscillators (F4-O1 and F3-O2); however, there was difficulty localizing the 5th uncoupled generator at Cz. The ICA and sLORETA methods did the best job at identifying the Cz generator, while the other methods are not close. Overall estimates of localization error for each method is given in Table 5.1.

For the ICA solution, localization was performed by fitting components as dipoles within the solution space using the dipFIT plug-in [208] to EEGLAB [71] shown in Figure 5.6. In addition to the localization parameters, each dipole also has three orientation parameters. The four coupled oscillators are oriented correctly, while the generator at Cz is oriented forward instead of upwards. This reinforces the fact that aspects of the third generator are lost during mixing due to volume conduction and the spatial properties of the model.

Method	MNE	sLORETA	LASSO	VB	ICA
MSE	0.2200	0.1517	0.3279	0.3138	0.1205

Table 5.1: Mean squared localization errors for the estimated sources by each method. The best localization results are from the ICA algorithm, followed closely by sLORETA and then MNE. All methods appear to do an acceptable job localizing the four coupled generators, but have difficulty with the uncoupled generator. The ICA method comes closest to identifying the generator at Cz, followed by sLORETA and MNE; both the VB and LASSO solutions completely fail to identify the 5th source, and have the highest localization errors.

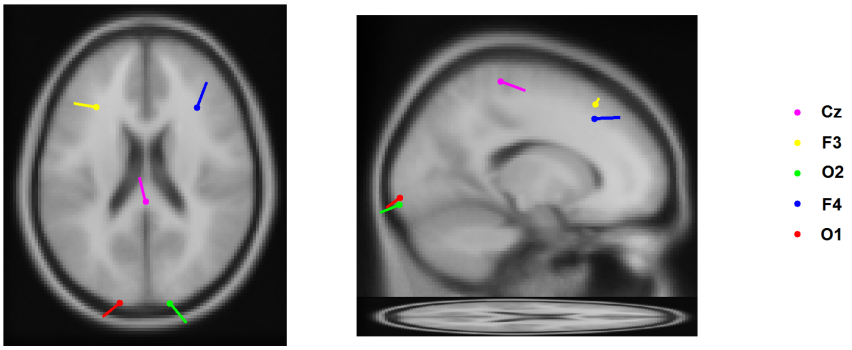

Figure 5.6: Top down view (left) and side view (right) of fitted dipoles of ICA components calculated and plot using the dipFit plug-in [208] to the EEGLAB toolbox [71] for Matlab. Each generator is correctly localized, with both pairs of couple oscillators oriented in approximately the correct directions and the uncoupled oscillator at Cz oriented in the wrong orientation.

5.3.4 Phase Shift Results

To assess each methods ability to correctly localize phase shift events, the cumulative summation (CUSUM) and phase derivative (PD) algorithms are applied to selected pairs of reconstructed sources. The above simulation experiment is replicated 12 times for a total of 111 shift events for this analysis. Based on a visual inspection of signal PSDs,

each method is able to clearly identify the two pairs of coupled oscillators to be subjected to phase shift analysis. The results of the shift identification for each pair of generators are combined and used to generate receiver-operator-characteristic (ROC) curves for each source reconstruction technique and each measure, and the resulting curves are shown in Figure 5.7. The CUSUM estimator outperforms the PD estimator for each source reconstruction technique, this result is expected based on the results of applying the estimators directly to Rössler data in Chapter 3. The maximum accuracy (mACC) and area under ROC (AUROC) values are given in Table 5.2 for the CUSUM estimator and Table 5.3 for the PD estimator. Among the CUSUM estimators, the MNE and sLORETA source reconstruction techniques performed best, with ICA and VB not far behind. For the PD estimator, only the MNE and sLORETA solution performed better than chance (AUROC ¿ 0.5), while the VB, LASSO and ICA all had maximum accuracy values of 0.5. Overall, the MNE solution with CUSUM estimator performed best, with maximum accuracy of 0.83 and AUROC values of 0.841.

5.4 Conclusions

I have presented a simulation study to assess several common source connectivity methods in their ability to accurately estimate phase shift events both temporally and spatially. The MNE algorithm had the best results, although they were comparable to accuracy of the sLORETA, ICA and VB solutions. Among the distributed sources, the sLORETA algorithm has the best localization. The key difference between previous studies and what I have presented here is the precise localization of temporal events, as opposed to a overall estimate of connectivity over time. The former makes use of the superior temporal resolution of EEG as compared to fMRI, and facilitates an analysis of the balance between synchronization and desynchronization dynamics.

The CUSUM identification method clearly outperforms the PD solution in this situation. This is consistent with the results from Chapter 3, where the CUSUM method worked

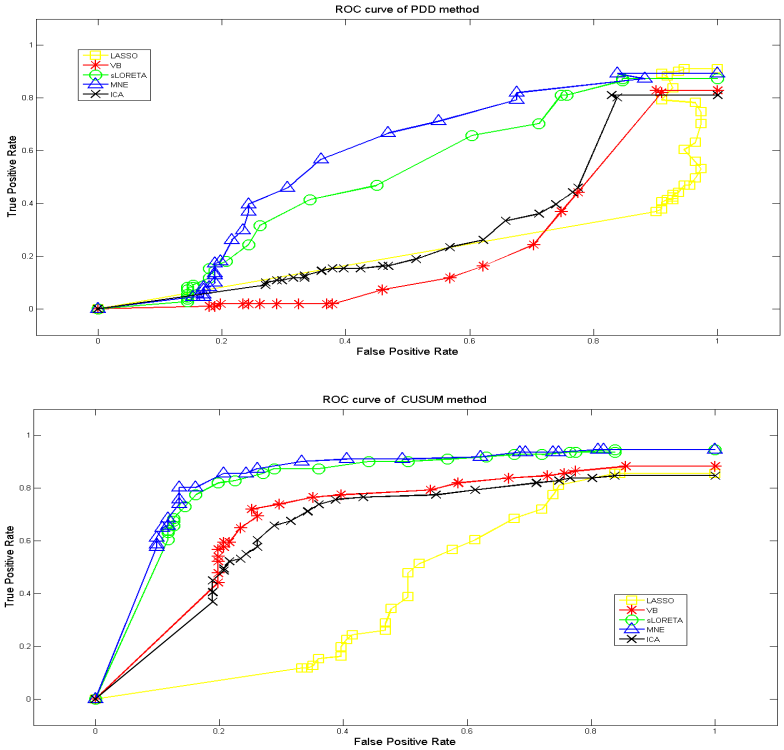

Figure 5.7: ROC curves for the PD method (top) and CUSUM method (bottom) of shift identification. The CUSUM method outperforms the PD method across the board, having greater accuracy and AUROC for every method. This is consistent with the results of applying these methods directly to Rössler data in Chapter 3. The PD method is able to identify the shift events; however, it is unable to avoid also false positives, due to the spatial and temporal nature of system noise. The CUSUM method performs much stronger, and is able to identify all of the shift events while keeping false positives to a minimum. The MNE and sLO solutions have the highest possible accuracy, with ICA and VB slightly behind. The maximum accuracy and AUROC values for each source reconstruction technique are shown in Table 5.2 for PD estimator and 5.3 for CUSUM estimator.

Method	MNE	sLORETA	LASSO	VB	ICA
AUROC	0.841	0.825	0.425	0.692	0.656
Max Accuracy (α)	0.83 (0.0001)	0.81 (0.0001)	0.53 (0.2154)	0.73 (0.0002)	0.69 (0.0022)

Table 5.2: The maximum accuracy and AUROC values for the CUSUM method. The four methods which perform well (MNE, sLORETA, VB and ICA) all have maximum accuracy values of at least 0.69 at significance levels around $\alpha = 0.0001$, while the LASSO method reaches a maximum accuracy of 5.3 at $\alpha = 0.2154$. The MNE solution has the maximum AUROC value at 0.841 and maximum accuracy of 0.83.

Method	MNE	sLORETA	LASSO	VB	ICA
AUROC	0.572	0.504	0.231	0.227	0.299
Max Accuracy (α)	0.60 (0.0464)	0.54 (0.0215)	0.5 (0)	0.5 (0)	0.5 (0)

Table 5.3: The maximum accuracy and AUROC values for the PD method. None of the methods perform especially well using the PD estimator, and only the MNE and sLO solution performs better than chance ($AUROC = 0.572, 0.504 > 0.5$), the other four methods have AUROC values below 0.5 which is undesirable. The MNE solution has the maximum accuracy value at 0.6. Both the LASSO, VB and ICA solutions have maximum accuracy of 0.5, meaning there is no value of α where they perform better than chance.

when applied directly to the Rössler attractors. The problem with the PD solution was an inability to avoid false positives; the use of SR techniques introduced noise into the signal which affects the estimation of the phase derivative, resulting in many spurious peaks in the PD. The temporal resolution of the PD method is not important for the Rössler application, but may still be necessary for other (EEG) applications. Furthermore, if the goal is to post-process these phase shift events by modeling them with the PSGC measure, then these additional false positives may not be a problem. We have seen in Chapter 4 that a less conservative threshold criteria increased the power of PSGC, and did not result in additional false positives in connectivity. When the false positives from the PD do not have any temporal structure, it may be more important to find all true shifts rather than minimizing false positives.

The LASSO solutions has the worst results of the different solutions; however, this may be due to the particular implementation of sparsity. As an example, sparsity on all parameters encourages sources in a specific orientation (X, Y or Z) by penalizing each parameter needed to describe the orientation. An alternative would be to have a sparsity penalty on the number of sources, but then fit the orientation and time course without sparsity penalty.

With the exception of ICA, each of the source reconstruction techniques presented assumes independence between time points. One way to boost the results could be to use more complex models with temporal structure, although this would come with increased computational requirements for the solutions. The VB solutions is the slowest of the current algorithms, but the Bayesian hierarchical model has a modular design which will allow different spatial and temporal constraints to be efficiently implemented and compared. Another options is a modification of the VB algorithm to simultaneously estimate shift and location, in the spirit of DCM source connectivity techniques.

5.4.1 Future Work and Limitations

One aspect of this work that needs to be addressed before these methods could be practically applied to EEG recordings is an automated method of ROI selection. Another investigation which could be undertaken is to give a specific description of geometry of sources which can be properly recovered and localized. This could include an investigation into the effect of additional interference (systematic, or spatial noise), or tangentially oriented sources on the results of localization.

In the final chapter, I give summary of the main results and impact of the thesis and provide concluding thoughts and future directions ideas for phase shift based measures of connectivity in EEG.

Chapter 6

Conclusions and Future Work

In this thesis, three different aspects of EEG phase analysis were considered; estimation of phase shift events, estimation of neural connectivity from phase shift events, and source reconstruction of phase shift events. These problems are considered within the context of EEG recordings, and there are several considerations which must be made for these analyses. Methods must be robust in the presence of high levels of noise. Recordings are subject to *volume conduction*, a linear mixing of generating signals. Methods should make use of the high temporal resolution of EEG recordings, one of the strengths of EEG recordings over more high spatial resolution techniques such as fMRI.

The primary impact of this thesis is in neuroscience, from the introduction of a novel method for estimating neural connectivity from the phase dynamics of EEG recordings. Each Chapter considers a different aspect of estimating neural connectivity from EEG phase dynamics, and overall they form a complete package for the estimation and interpretation of connectivity from EEG phase shift events. There are two other primary areas of impact; in change-point analysis of complex systems, and in analyzing EEG recordings in behavioural psychology [167]).

The synchronization and spontaneous desynchronization of neural populations is considered to be functionally relevant to neural processing, and an understanding of these

dynamics may provide valuable insight into brain dynamics. As we explore the data from many modes brain imaging techniques, we develop better theories of how these dynamics are generated. There are several theories which attempt to explain the existence of heavy-tailed power-law distributions; examples of this are self-organized criticality (SOC) [19], generalized q-entropy [284] or the free energy principal [96]. The existence of power-law behaviour occur often in neuroscience, has been shown to be functionally/behaviourally relevant, and so it is beneficial to understand how they are generated.

Previous studies involving phase shift identification in the EEG literature use a subjective threshold procedure [37, 277]. This reduces the reliability of the results as well as the reproducibility of EEG studies performing such analyses. The CUSUM and PD estimators provide an objective means of identifying phase shift events. A more general application exists in complex dynamical systems, where these results provide the tools for phase shift identification from noisy observed data.

It has already been seen that phase shift dynamics are related to developmental changes [277] and behavioural measures of anxiety or aggression [167]; the result of the simulation studies provide additional insight into the interpretation of EEG phase shift analyses, for our PSGC measure, but also in phase reset analysis [277]. As sophistication of analysis method (source reconstruction, connectivity analysis) and recording techniques (joint EEG-fMRI) continues to increase, this will result in a increased understanding about the role of synchronization in behaviour.

6.1 Future Work

Chapter 2 introduces a novel measure of estimating neural connectivity, phase shift Granger causality (PSGC). Future work on this measure will involve creating a better model for phase shift events, perhaps including a refractory period, or allowing for correlations between shift events.

In Chapter 3, two methods are explored for the identification of phase shift events, the cumulative summation (CUSUM) and phase derivative (PD) estimators. A future goal of this work is to modify the proposed methods to work in real-time, allowing immediate feedback to be provided to the user. The PD methods seems most natural for such an application, with high temporal resolution and computationally efficient, but perhaps a modification of the CUSUM estimator could be found to help it adapt to real-time computation. Another potential extension of this work is to use bootstrapping techniques to estimate the variability in phase shift localization, and use this information to construct a confidence interval on the true location of the shift event [142].

Finally, I explore the effect of volume conduction on phase shift identification, and the potential to use source reconstruction events to nullify this effect, while concurrently increasing the spatial resolution of the EEG analysis. In Chapter 4 it is determined that the PSGC measure is provides an appropriate measure of neural connectivity under certain assumptions on the underlying sources; they should be oriented perpendicular to the surface of the brain, and there must be enough spatial noise to attenuate the spread of activity due to volume conduction. In Chapter 5 several methods of EEG source reconstruction are applied to determine which method is best maintains the phase shifting dynamic in the original signal, and determined that the minimum norm estimate (MNE) works best. Future work on this topic will be to extend the PSGC measure to work on reconstructed sources, rather than on scalp recordings directly.

APPENDICES

Appendix A

Forward Problem

The forward problem can be defined as determining the electric field potential that is generated by a given primary current density. Specifically, interest lies in solving for the potential at the scalp due to primary currents within the brain. The equations which describe the behaviour of electricity and magnetism are called Maxwell's equations,

$$\begin{aligned} \nabla \cdot E &= \frac{\rho}{\epsilon_0} \\ \nabla \cdot B &= 0 \\ \nabla \times E &= -\frac{\partial B}{\partial t} \\ \nabla \times B &= \mu_0 J + \mu_0 \epsilon_0 \frac{\partial E}{\partial t}, \end{aligned}$$

where E is the electric field, B is the magnetic field, ρ the total electric charge density, ϵ_0 the permittivity of free space and μ_0 the permeability of free space. In the context of EEG recording the contribution of the time derivatives is negligible [129]; the quasi-static approximation to Maxwell's equations is used, setting the time derivative terms $\frac{\partial B}{\partial t}$ and $\frac{\partial E}{\partial t}$ to zero. A consequence of this approximation is $\nabla \times E = 0$ which implies that there exists a scalar field potential (Y) such that $E = -\nabla Y$.

Using Maxwell's equations, the relationship between the electric potential Y and the primary current density J_p [104, 129]

$$\nabla^2 Y = \frac{\nabla J^p}{\sigma}. \tag{A.1}$$

To see this, first separate the electric current density into the contributions from the primary current source J^p and the secondary currents due to the electric field, J^s

$$J = J^p + J^s.$$

The secondary current density of a biological system is linearly related to the electric field for low voltage primary currents (below a threshold of about 2 milliAmp/cm-squared) such as those generated in the brain/heart [251],

$$J^s = \sigma E,$$

where σ is a conductivity parameter. This makes the total current density

$$J = J^p + \sigma E.$$

Substituting $E = -\nabla Y$

$$J = J^p - \sigma \nabla Y,$$

and then taking divergence

$$\nabla J = \nabla J^p - \sigma \nabla^2 Y = 0.$$

By conservation of energy principal, the divergence of the total electrical current is zero, $\nabla \cdot J = 0$, so

$$\nabla^2 Y = \frac{\nabla J^p}{\sigma}$$

In the case of a homogeneous volume, the following solution to Equation A.1 is [22, 129],

$$Y(r) = \frac{1}{4\pi\sigma} \int \frac{J^p \cdot d}{\|d\|^3} dv + \frac{1}{4\pi} \int_{\partial V} Y \cdot \hat{n} \frac{d}{||d||^3} ds. \tag{A.2}$$

This equation can then be solved either analytically (for simple volumes) or numerically to determine the electric potential Y.

Appendix B

Filtering

B.1 Noise

The nature of EEG recordings is such that there is the potential for high levels of noise, there are some standard considerations before attempting to meaningfully analyze a recording.

In EEG recordings an artefact is any contribution to the field potential that is not due to brain activation. Artefact detection and removal is necessary for many types of EEG analysis [58]. Some artefacts enter the signal in a systematic way, allowing the original signal to be recovered through using signal processing techniques. An example is the artefact generated by the electrical equipment that is near the recording area, this manifests as a strong 60Hz component in the signal (in North America). The electric potential is also affected during muscle activity, an eye blink has a very distinct electrical signature (see Figure 1.10). Depending on the nature of the experiment, the effect of blink artefacts may be removed using regression techniques [174, 297]. Other muscle activity such as such as jaw clenching or frowning is less well behaved and most data which contain such activity must be omitted from analyses.

One type of preprocessing used in EEG analysis is *segmentation* , taking a continuous EEG recording is dividing it into many smaller *epochs*. The main reason for segmenting a dataset is to create sets of data which are approximately stationary. This is important for statistical analyses, as EEG in general is non-stationary beyond a few seconds [55, 75]. When using segmentation, epochs with extreme noise can be easily omitted from analyses without affecting the rest of the data. In the study of event related potentials (see Section 1.4.2), segments are often time locked to a specific event. By having each segment start just prior to the onset of the event, the properties of each waveform are more likely to capture the relevant brain activity.

Another stage of preprocessing is to select a reference for the data. A property of field potentials is that they can only be measured up to a constant, meaning that one must look at the difference between electric potentials in order to generate a physically meaningful quantity [58]. In EEG recordings differences are obtained by comparing the electric potential at each electrode to a common source called the reference electrode. The ideal reference point is one where there is no electrical activity present however there is no place on the scalp which fits this criteria. Reference electrodes are often placed on the mastoid bones just behind the ears, because they are thought to be the area with the lowest amount of electrical activity. Another option is to use the difference between adjacent electrodes to create a bipolar reference. The advantage of a bipolar reference is that it is able to capture localized activity, however global activity that is common to both electrodes is lost during this process. Once a recording has been referenced to a common electrode, one option is to rereference against the average voltage of all the electrodes, this method is independent of the initial reference point. Each of these options has strengths and weaknesses and it is up to an experienced researcher to determine which is best for any application [125, 229].

B.2 Filtering

One technique for artifact removal is to use a *filter*, an operator designed to remove unwanted spectral components of a signal. If the frequency of the artifact component can be identified, a notch filter can be used to remove contributions from the exact frequency. In other situations a bandpass filter can be used, these filters remove all contributions from frequencies outside the designated frequency band. This method is most useful when the artefacts occur outside the frequency band of interest. When using filtering in broader analyses, there is always the possibility of removing important signal content.

One type of noise which is often dealt with using filters is the contribution due to electrical equipment in the recording room, in North America this manifests as a 60Hz oscillation. A one second waveform from location C3 is plotted along with its corresponding PSD in Figure B.1, notice the spike at 60Hz. One option to remove this artifact is to use a notch filter, to removing only the information at 60Hz. Often analyses are only concerned with signals in a specific frequency band, so instead the data may be run through a 30Hz low pass filter to remove all contributions above 30Hz including the 60Hz noise.

B.3 Filters and Phase

An important aspect of the complex demodulation calculations is the restriction that there only be a single frequency component in the signal. To achieve this, the signals are put through filters which remove the unwanted components of the signal. The ideal filter which selects specific frequency components does not exist, so we must chose some approximation of this. There are many different filter designs which can be used, here we investigate the effect of using different filters on phase shift identification and latency estimation.

The two most common classes of digital filters are the finite impulse response (FIR) and the infinite impulse response (IIR). Using FIR filters requires more data and more

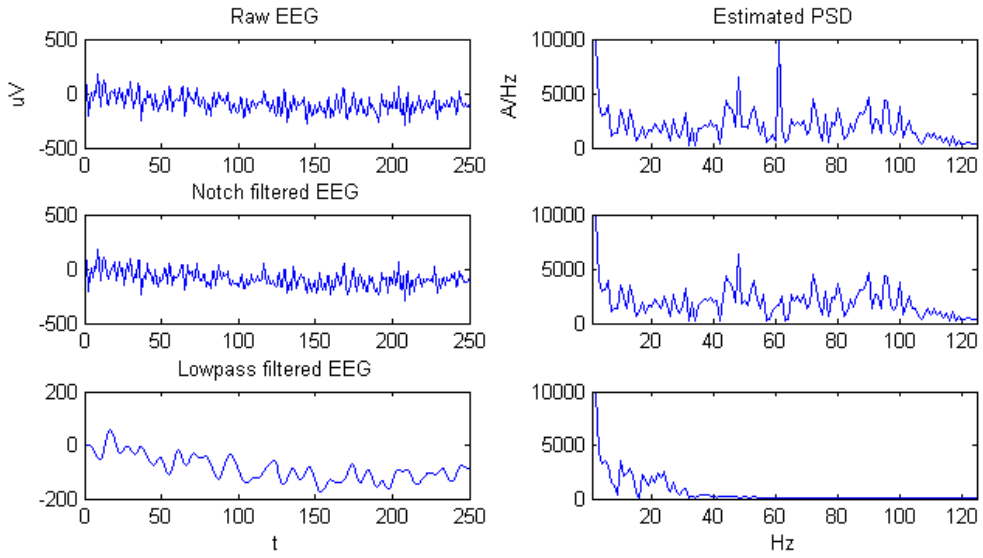

Figure B.1: Location C3, segment 3. Top: Raw data and power spectrum. Middle: After 60Hz notch filter. Bottom: After 50Hz low pass filter.

computation time than IIR filters, however they are always stable. In contrast, IIR filters are much faster and are able to use lower order filter to achieve similar accuracy, however they maybe become unstable.

There are four classical IIR filter designs, the Butterworth, Chebychev1, Chebychev2 and Elliptic filters. Aside from the order of the filter, there are two possible parameters of interest in filter design, pass-band ripple and stop-band ripple. The Butterworth design does not constrain either parameter, while the Elliptic constrains both of them. The Chebychev1 design constrains only the pass-band ripple, while the Chebychev2 design constrains the stop-band ripple.

The two filters which constrain the stop-band ripple (Chebychev2 and Elliptic) are sensitive to the order of the filter, and for they may be unstable. Depending on the order, sometimes they introduce high frequency spiking into the instantaneous phase which pre-

vents the identification of multiple shifts. A sample of this instability is shown in Figure B.2, where the instantaneous phase of a signal is shown as estimated by a stable 4th order Butterworth filter and an unstable chebychev2 filter of the same order. The signal has two phase shift events each with a magnitude of $3\pi/4$ and which occur at 1000 and 1700 ms.

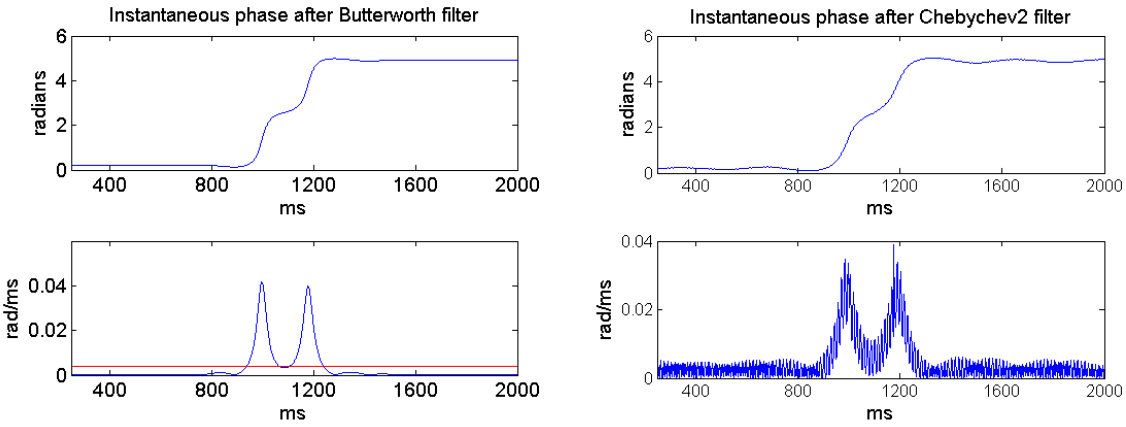

Figure B.2: Left: 4th order Butterworth filter. Right: 4th order chebychev2 filter with 20db of stop-band ripple. The Butterworth filter can easily identify the two shifting events, while the Chebychev2 filter introduces a high frequency ripple into the phase of the signal which makes identification using the threshold method impossible.

The next step will be to evaluate the two stable filters on their ability to accurately identify phase shifts. To do this I will look at the minimum ISI in the noise-free paradigm and the mean and standard deviation of the latency error in the $(SNR = 1, \Delta\phi = 3\pi/4)$ paradigm. Figure B.3 shows the minimum ISI δ such that both shifts can be identified as a function of filter order; the filters used are the Butterworth filter and three Chebychev1 filters with maximum pass-band ripple $R = [0.1, 0.5, 1]$. The results show that the minimum ISI is a consistent 300-350 ms for the Butterworth filter, while the Chebychev1 filters (R=0.1 or 0.5) have results between 100-200 ms for low filter orders and between 200-300 ms for higher orders. The results of the latency errors shows that the mean latency error is consistently around 3-4 ms for all filters and orders. The standard deviations plateau

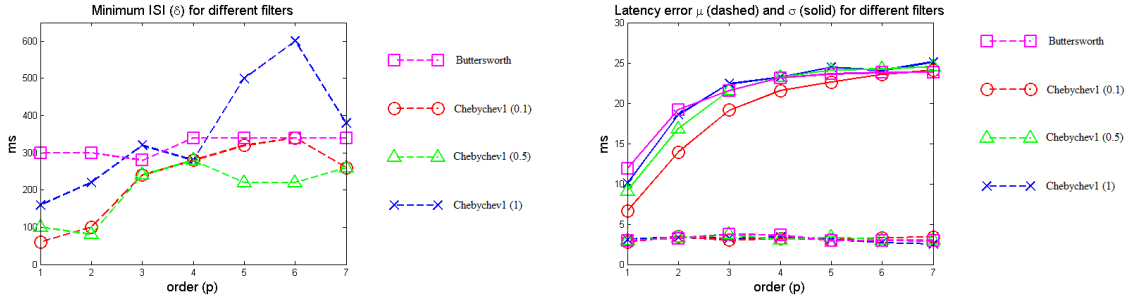

Figure B.3: Plots of the behaviour of four different filters for different orders; there is a Butterworth filter and three Chebychev1 filters with pass-band ripple R=[0.1, 0.5, 1]. Left:

with increasing filter order, ending up in the 20-25ms range.

Appendix C

Variational Bayes

The notation for the Bayesian analysis will be that $f(y, \theta)$ is the joint distribution between the vector of observed data and the vector of parameters, $p(\theta)$ the prior distribution, $p(\theta|y)$ the posterior distribution, m(y) the marginal distribution and $f(y|\theta)$ the likelihood function.

$$p(\theta|y) = \frac{f(y,\theta)}{m(y)} = \frac{f(y|\theta)p(\theta)}{m(y)}$$

C.1 Variational Bayes

Variational Bayes (VB) is a parametric method for approximating the posterior distribution in Bayesian analysis. The method is typically applied in situations where the complex nature of the problem makes analytic solutions difficult or impossible. In VB, the attention is to restrict attention to a set of candidate distributions which factor into groups of parameters, and from these candidates selecting the approximate posterior which maximizes the *variational free energy* (VFE). With an appropriate factorization of parameters, this often leads to an analytically tractable and efficiently computed solution. The method was first applied [87] for evaluating line integrals, it later appears in the statistical literature in the context of Bayesian inference for complex models [12, 231].

Model Evidence

The KL divergence (see E) from the true posterior to the approximation is one way to assess the accuracy of the approximation, however since the true posterior is unknown this quantity cannot be used directly. The VFE provides an alternative way to assess the quality of an approximation.

Definition C.1. *In Bayesian inference, the variational free energy F(q) of an approximate distribution $q(\theta)$ is defined as*

$$F(q) = \int q(\theta) \log \frac{f(y,\theta)}{q(\theta)} d\theta,$$

where $f(Y,\theta)$ is the joint distribution of the data and the parameters.

To see the connection between VFE and the KL divergence, consider the following decomposition of the log model evidence,

$$\begin{aligned} \log m(y) &= \int q(\theta) \log m(y) d\theta \\ &= \int q(\theta) \log \frac{f(y,\theta)}{q(\theta)} d\theta + \int q(\theta) \log \frac{p(\theta \mid y)}{q(\theta)} d\theta \\ &= F(q) + KL(p|q) \end{aligned}$$

The significance of this result is that since $\log(m(y))$ does not depend on the approximate posterior, maximizing the VFE is equivalent to minimizing the KL divergence. The VFE can be decomposed into two parts, the first is the expected log likelihood of the data under the approximate posterior and the second is the KL divergence of the prior distribution from the approximate posterior.

$$\begin{aligned} F(q) &= \int q(\theta) \log \frac{f(y,\theta)}{q(\theta)} d\theta \\ &= \int q(\theta) \log f(y \mid \theta) d\theta - \int q(\theta) \log \frac{q(\theta)}{p(\theta)} d\theta \\ &= E_q(\log L(\theta)) - KL(q \mid p) \end{aligned}$$

Mean Field Approximation

For a partition Θ of the parameters, a mean field approximation of the distribution $p(\theta)$ is a distribution $q(\theta)$ such that elements of Θ are independent.

Definition C.2. *If $\Theta = \{\theta_1, \theta_2, ..., \theta_k\}$ is a partition of θ, then a mean field approximation of a distribution $p(\theta)$ over Θ is any probability distribution with the form*

$$q(\theta) = \prod_{i=1}^{k} q_i(\theta_i).$$

A mean field approximation removes the posterior covariance between elements of the Θ. Although the stochastic dependence between variables is removed, there can still be systematic dependence based on the average effect of other variables.

In VB, the form of the distributions $q_i(\theta_i)$ which maximize the VFE of a made field approximation can be found using the following results [231],

Theorem C.1. *If $\Theta = \{\theta_1, \theta_2, ..., \theta_k\}$ is a partition of θ and $q = \prod q_i(\theta_i)$ is an approximation of $p(\theta)$ then the form of q_i which maximizes the VFE of $q(\theta)$ is*

$$q_i(\theta_i) = \frac{e^{I(\theta_i)}}{C_i}$$

where

$$I(\theta_i) = \int q(\theta_{-i}) \log\left(L(Y \mid \theta) p(\theta)\right) d\theta_{-i}, \quad \theta_{-i} = \{\theta_j \in \Theta | j \neq i\}. \tag{C.1}$$

This theorem can be used to identify which exponential family the approximate posterior $q(\theta_i)$ belongs to. The natural parameters for the distribution are calculated using the function $I(\theta_i)$, which is the expected log-posterior distribution with expectation of all other variables taken with respect to the approximate posterior. This results in a set of natural parameters for $q(\theta_i)$ which are a function of the expectations of other variables $E_q(\theta_{-i})$.

Estimation

The *Expectation-Maximization* (EM) algorithm is an iterative method often used for estimating parameters in the presence of missing data or hidden variables [74]. An adaptation of the EM algorithm can be used to find the approximate posterior which maximizes the VFE [244]. The maximization step uses current values of the parameters to find the form of the approximate posterior which maximizes the VFE, then the expectation step updates the parameter values using the expectation of the approximate posterior. Starting with an initial value for each parameter, the estimates for each group of parameters are iteratively updated until they have converged to a (potentially local) maximum of the VFE.

- Start with an initial values for each parameter $\theta^{(0)}$
- Calculate the VFE, $F(q^{(0)})$
    1. For each θ_i, calculate new values $\theta_i^{(t)}$ based on previous values $\theta_{-i}^{(t-1)}$
    2. Calculate $F(q^{(t)})$ for new values of parameters
    3. If $\frac{\left|F(q^{(t)})-F(q^{(t-1)})\right|}{F(q^{(t-1)})} < \epsilon$ then stop, otherwise go to 1.

Example 1

As a simple example, consider a dataset $\{y_i\}$ of n independent observations from $Y \sim N(\mu, \lambda)$, where $\lambda = \frac{1}{\sigma^2}$ is a precision variable. The prior distribution is $\mu \sim N(0, 1)$ independent of $\lambda \sim Ga(\frac{1}{2}, 5)$. This leads to a posterior distribution of the form

$$p(\mu, \lambda|y) = \frac{1}{m(y)}\lambda^{n+1}e^{\frac{-\mu^2-4\lambda-\lambda\sum_i(y_i-\mu)^2}{2}}.$$

This particular posterior distribution is not complicated and can be solved analytically, this will allow us to investigate the accuracy of the VB approximation. The grouping of parameters is trivial in this example, there are only two parameters. To find the optimal

form of $q(\mu)$, use Equation C.1

$$\begin{aligned} I(\mu) &= \int q(\lambda) \log(L(\mu,\lambda)p(\mu,\lambda))d\lambda \\ &\propto \int q(\lambda) \left(\frac{-\mu^2 - \lambda\mu^2 + 2\lambda n\bar{y}\mu}{2} \right) d\lambda \\ &= -\frac{\mu^2}{2}(1 + nE_q(\lambda)) + \mu(n\bar{y}E_q(\lambda)) \end{aligned}$$

Comparing this to a normal density leads to the optimal form of $q(\mu)$,

$$q(\mu) \sim N\left(\frac{n\bar{y}E_q(\lambda)}{1 + nE_q(\lambda)}, \frac{1}{1 + nE_q(\lambda)} \right).$$

By a similar calculation, the optimal form of $q(\lambda)$ is given by,

$$q(\lambda) \sim Ga\left(\frac{1}{2} + n, \left[\frac{1}{5} + \frac{\sum_i y_i^2 - 2n\bar{y}E_q(\mu) + nE_q(\mu^2)}{2} \right]^{-1} \right).$$

Data for this experiment was simulated from a N(-0.3, 1.5) distribution. The initial values that were used are $(\mu_0, \lambda_0) = (0, 1)$ and the convergence criterion $\epsilon = 10^{-5}$. The experiment was repeated for several values of n, and the results recorded in Table C.1. The result of the simulation is that the VB approximation appears to be converging to the true posterior distribution, this is to be expected as the parameters are independent in the prior.

n	iterations	$E_q(\mu)$	$E_\Pi(\mu)$	$E_q(\lambda)$	$E_\Pi(\lambda)$
10	5	0.0949	0.0943	1.8915	1.8920
25	4	-0.5506	-0.5498	1.2436	1.2436
50	5	-0.2656	-0.2655	2.7026	2.7026
100	4	-0.2811	-0.2810	5.9862	5.9859

Table C.1: Results of VB Example 1. The mean of the approximate posterior are converging to the mean of the true posterior. This is to be expected since the parameters priors are independent.

Example 2

Next consider an example where the prior distributions of μ and λ are not independent. In this situation the posterior distribution does not factor by groups of parameters, it will be interesting to see how this affects the accuracy of the VB method. To introduce a dependency in the prior distribution the distribution of λ will be conditional on the value of μ. For this prior μ is assumed to be distributed as N(0,1) and then $\lambda \mid \mu$ is $Ga(5, \frac{1}{\mu^2})$.

The use of a non-conjugate prior means that there is not necessarily a tractable solution to the problem, however the updates can still be easily computed numerically. The results of the second example can be seen in Table C.2. The mean of the approximate posterior is close to the mean of the true posterior, however the estimates are not as good the independent case. The variance of the approximation is an order of magnitude smaller than the true variance of the distribution; the VB algorithm does not provide a good estimate of variance in the case of posterior dependence between parameters.

n	$E_q(\mu)$	$E_\Pi(\mu)$	$E_q(\lambda)$	$E_\Pi(\lambda)$	$V_q(\mu)$	$V_\Pi(\mu)$	$V_q(\lambda)$	$V_\Pi(\lambda)$	ρ_Π
20	2.7857	2.2330	0.1859	0.1488	0.1550	1.3697	0.0013	0.0066	0.8592
50	3.6007	1.7652	0.1491	0.0731	0.1047	3.2900	0.0004	0.0057	0.9755

Table C.2: Results of VB example 2. The parameters used to generate the model are $\mu_0 = 1$ and $\lambda_0 = 2$. The mean of the approximate posterior is close to the mean of the true posterior, however the estimates are not as good the independent case. The variance of the approximation is an order of magnitude smaller than the true variance of the distribution. The VB algorithm does not provide a good estimate of variance in the case of dependence.

C.2 EEG Hierarchical Regression Model

An example of using the VB method to solve the source reconstruction problem using the 3-tier Hierarchical regression model used in [282], which is described graphically in Figure

C.1. The response variable $Y_{s\times t}$ is a matrix of electric potentials recorded from s sensors at t time points. The top level of the model relates Y to $J_{g\times t}$, the matrix of primary current densities from g generators at t time points.

$$Y|J = KJ + \epsilon_Y$$

The known matrix $K_{s\times g}$ is the lead field matrix (LFM) which solves the EEG forward problem (see Section 1.6.1), the term ϵ_Y represent the measurement errors at each sensor. These errors are independent of each other and the errors at other level, the errors for each sensor have a different precision determined by the vector $\sigma_{s\times 1}$.

In the second level of the model, a temporal regression is used to describe the time evolution of the primary current density in terms of a matrix of regression coefficients $B_{k\times g}$ for k regressors at g generators

$$J^T|B = XB + \epsilon_J.$$

The design matrix $X_{t\times k}$ contains the value of k predetermined temporal regressors evaluated at t time points (see Section C.3). The term ϵ_J represents the lack of fit for the temporal model, these errors are independent of all other errors and generator has a precision defined by the vector $\lambda_{g\times 1}$.

The final level of the model describes the spatial structure of the coefficient matrix B,

$$B = \epsilon_B$$

The coefficients are modelled as having zero mean and the term ϵ_B represents the lack of fit of the spatial model. Spatial information in this level of the model can be included in the prior covariance structure of ϵ_B. The errors ϵ_B are independent from errors on at other levels of the model and between regressors, but the spatial information in the model is contained in the covariance matrix between generators $D_{g\times g}$.

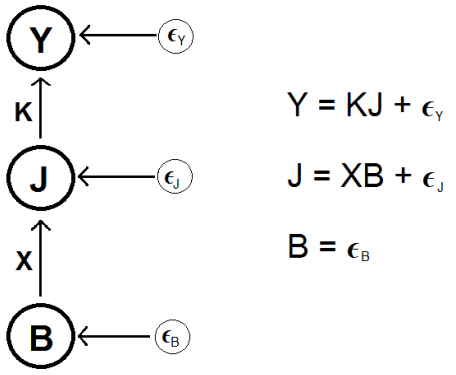

Figure C.1: A graphical representation of the hierarchical linear model used for the EEG inverse problem. Circular terms represent additive error terms at each level of the model. The red boxes represent the unknown parameters in the model, while the other boxes are known.

C.2.1 Prior Distribution

When dealing with matrix random variables, the Kronecker product is used to describe the covariance of the distribution. The first argument defines the covariance in a column of the matrix, while the second argument is the covariance in a row of the matrix. This is a slight abuse of notation, as the Kronecker product actually defines the covariance of the vectorized matrix.

The error terms in the hierarchical model are assumed to follow a multivariate Gaussian distribution with mean zero. For the top level of the model, the errors are independent across time and sensor, the vector σ is a set of parameters which allows the precision of each sensor to be different

$$\epsilon_Y \sim G\left(0, [\sigma I_s \otimes I_t]^{-1}\right).$$

The errors in the second level are independent across time and generator, the vector λ allows the precision to be different at each generator

$$\epsilon_J \sim G\left(0, [\lambda I_g \otimes I_t]^{-1}\right).$$

At the third level, the errors are independent across temporal regressors with the precision of each regressors defined by the vector α. The matrix D defines the covariance between the generators, the form of D is chosen to represent prior information about the spatial characteristics of the solution (see Section C.3.1),

$$\epsilon_B \sim G\left(0, [D \otimes \alpha I_k]^{-1}\right).$$

The prior distribution for the precision hyper parameters will use independent gamma distribution, each with their own hyper parameters,

$$p(\alpha) = \prod_{i=1}^{k} Ga(a_{\alpha,i}, b_{\alpha,i}).$$

This results in conjugate normal-gamma prior distributions for mean-precision parameters. The use of gamma priors for the precision variables is also called *automatic relevance detection* (ARD) priors [198], ARD priors allow the data to determine which regressors, generators and sensors (corresponding to α, λ, σ) are most relevant in the model.

C.3 Experiment

In this experiment there are $g = 39$ generators located in $\mathbb{R}^3$. The generators are arranged as the surface of a 5 by 3 by 3 box centred at (0,0,0), except there are no generators at the bottom of the box. The scalp is the surface of a sphere centred at (0,0,0) with radius 5. There are $S = 4$ sensors on the scalp, with one in each quadrant of the upper hemisphere. A visualization of the head model can be seen in Figure C.2.

The LFM used in the model will be that of a spherical head model with homogeneous isotropic conductivity, and in addition the primary current density is constrained to be oriented perpendicular to the surface. If r_i is the location of the i^{th} generator, s_j the location of the j^{th} sensor and $\sigma = 1$, the formula for the elements of K is [192]

$$k_{i,j} = \frac{1}{4\pi}\left(\frac{2(r_i \cdot s_j) - \|s_j\|\,\|r_i\|}{\|r_i - s_j\|^3} + \frac{1}{\|r_i\|\,\|r_i - s_j\|} - \frac{1}{\|r_i\|\,\|s_j\|}\right)$$

The temporal design matrix X contains the values of $k = 4$ temporal regressors evaluated at each time point. The four temporal regressors are $\sin 2t$, $\sin 3t$, $2^{-t/2} \sin 5t$ and t. The matrix D represents prior knowledge about the spatial dependence of B, for this simple example it is just the identity matrix $D = I_g$.

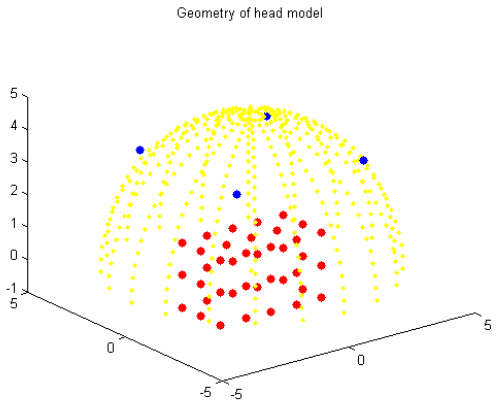

Figure C.2: A visualization of the geometry for the head model used in the simulation. Red dots represent the source generators arranged as a bottomless box in the centre of the head. Yellow dots represent the scalp of the head, while blue dots represent EEG electrodes which record the signal.

Following [282], the values of the hyper parameters in the ARD priors were taken to be $a = \frac{1}{1000}$ and $b = 1000$. This results in an uninformative prior distribution which has expected value of 1 and a variance of 1000, allowing the data to dominate the estimates of precision parameters. The effect of different values of the hyper parameters is explored in [203].

C.3.1 Simulated Data

To simulate the observed data for this experiment, a set of regression coefficients is selected to represent the source activity. The source is centred at (2,0,1) where the coefficient is

set to 4, generators which are one space away from the source have coefficients set to 2, generators two spaces away have coefficients of 1 and all other generators have coefficient 0.

To generate the primary current density from the regression coefficients, the coefficients are applied to two temporal regressors, $\sin(2t)$ and $\sin(3t)$. The primary current density is then randomly generated from a multivariate normal distribution with a variance of 0.01.

To generate the observed data Y, a small amount of jitter is added to the LFM,

$$k^*_{i,j} = k_{i,j} + N(0, 0.001).$$

The data is then generated using the solution to the forward problem,

$$y = K^* J.$$

C.3.2 Variational Bayes Solution

For the VB solution the approximate posterior is assumed to factor over the obvious groups of parameters, and the regression coefficients at each generator are also independent,

$$q(\sigma, \lambda, \alpha, J, B) = q(\sigma)q(\lambda)q(\alpha)q(J)\prod_{g=1}^{G} q(B_{g,\cdot}).$$

The decision to have coefficients at different generators be independent may seem counter-intuitive since the goal of these parameters is to add spatial information into the model; however, this information (the matrix D) is still a factor in the solution, through the mean of the posterior distribution for B.

It is not assumed explicitly in the VB solution, however in the model the columns of J are independent with equal covariance structure. Here are the update formulas for $q(J_{\cdot,t})$,

$$\begin{aligned} J_{\cdot,t} &\sim MVN(\mu_{J_t}, \Sigma_J) \\ \Sigma_J &= \left(K' diag(\sigma) K + diag(\lambda)\right)^{-1} \\ \mu_J &= \Sigma_J \left(K' diag(\sigma) Y_{\cdot,t} + diag(\lambda) B X_{\cdot,t}\right). \end{aligned}$$

The update formulas for $q(B_{g,\cdot})$,

$$\begin{aligned} B_{g,\cdot} &\sim MVN(\mu_{B_g}, \Sigma_{B_g}) \\ \Sigma_{B_g} &= (\lambda(g)XX' + diag(\vec{\alpha})D(g,g))^{-1} \\ \mu_{B_g} &= \Sigma_{B_g}\left(\lambda(g)XJ_{g,\cdot} - diag(\vec{\alpha})\vec{r}_g\right) \\ \vec{r}_g &= \sum_{i \neq g} D_{g,i}B'_{i,\cdot} \end{aligned}$$

It is not assumed explicitly, however the columns of J are independent with equal covariance structure. Here are the update formulas for $q(J_{\cdot,t})$,

$$\begin{aligned} J_{\cdot,t} &\sim MVN(\mu_{J_t}, \Sigma_J) \\ \Sigma_J &= \left(K'diag(\vec{\sigma})K + diag(\vec{\lambda})\right)^{-1} \\ \mu_J &= \Sigma_J\left(K'diag(\vec{\sigma})Y_{\cdot,t} + diag(\vec{\lambda})BX_{\cdot,t}\right) \end{aligned}$$

Update formulas for $\vec{\alpha}$, $\vec{\lambda}$ and $\vec{\sigma}$, the approximate posteriors are independent gamma functions each with their own hyper-parameters

$$\begin{aligned} a_{\alpha,i} &= a + \frac{G}{2} \\ b_{\alpha,i} &= \left(\frac{1}{b} + \frac{1}{2}B'_{\cdot,i}DB_{\cdot,i} + \frac{1}{2}\sum_{g=1}^{G} D_{g,g}[\Sigma_{B_g}]_{i,i}\right)^{-1} \\ \alpha_i &= a_{\alpha,i}b_{\alpha,i} \\ a_{\lambda,g} &= a + \frac{T}{2} \\ b_{\lambda,g} &= \left(\frac{1}{b} + \frac{T}{2}[\Sigma_J]_{g,g} + \frac{1}{2}\sum_{t=1}^{T}(J_{g,t} - B_{g,\cdot}X_{\cdot,t})^2 + X'_{\cdot,t}\Sigma_{B_g}X_{\cdot,t}\right)^{-1} \\ \lambda_g &= a_{\lambda,g}b_{\lambda,g} \\ a_{\sigma,s} &= a + \frac{T}{2} \\ b_{\sigma,s} &= \left(\frac{1}{b} + \frac{T}{2}K_{s,\cdot}\Sigma_J K'_{s,\cdot} + \frac{1}{2}\sum_{t=1}^{T}(Y_{s,t} - K_{s,\cdot}J_{\cdot,t})^2\right)^{-1} \\ \sigma_s &= a_{\sigma,s}b_{\sigma,s} \end{aligned}$$

For the VB solution, the initial conditions for the mean parameters μ_J and μ_B are set to 1, the initial covariance matrices Σ_J and Σ_{B_g} are the identity matrix, the precision hyper-parameters $b_\cdot$ and $a_\cdot$ are set to 1. The algorithm is run until the relative change in VFE is less than 10^{-6}, for this experiment it took 402 iterations and about 20 seconds to compute the solution. the trajectory of the VB solution of the parameters can be seen in Figure C.3. The ARD prior lets the data determine the relevance of each regressor, generator and sensor. Panels a) and c) show that there is one specific regressor/sensor which is most important in the solution, while panel b) shows a cluster of generators which are important. In panel f) it can be seen that the VFE is converging to a (potentially local) maximum.

C.3.3 MCMC Solution

Due to the linear hierarchical structure and the selection of conjugate prior distributions the full conditional distributions are available for this model. The full conditional distributions for each of the groups of parameters,

$$
\begin{aligned}
J_{\cdot,t} \mid \cdot &\sim MVN\left(\Sigma_J\left(K' diag(\vec{\sigma}) Y_{\cdot,t} + diag(\vec{\lambda}) B X_{\cdot,t}\right), \Sigma_J\right) \\
B_{g,\cdot} \mid \cdot &\sim MVN(\Sigma_{B_g}\left(\lambda_g X J'_{g,\cdot} + diag(\vec{\alpha}) R_g\right), \Sigma_{B_g}) \\
\alpha_p \mid \cdot &\sim Ga\left(a + \frac{G}{2}, \left(\frac{1}{b} + \frac{1}{2} B'_{\cdot,p} D B_{\cdot,p}\right)^{-1}\right) \\
\sigma_s \mid \cdot &\sim Ga\left(a + \frac{T}{2}, \left(\frac{1}{b} + \frac{1}{2}\sum_{t=1}^{T}(Y_{s,t} - K_{s,\cdot} J_{\cdot,t})^2\right)^{-1}\right) \\
\lambda_g \mid \cdot &\sim Ga\left(a + \frac{T}{2}, \left(\frac{1}{b} + \frac{1}{2}\sum_{t=1}^{T}(J_{g,t} - B_{g,\cdot} X_{\cdot,t})^2\right)^{-1}\right)
\end{aligned}
$$

$$
\begin{aligned}
\Sigma_J &= \left(K' diag(\vec{\sigma}) K + diag(\vec{\lambda})\right)^{-1} \\
\Sigma_{B_g} &= \left(\lambda_g X X' + diag(\vec{\alpha}) D_{g,g}\right)^{-1} \\
R_g &= \sum_{r \neq g} B'_{r,\cdot} D_{r,g}
\end{aligned}
$$

Comparing the full conditional distributions to the optimal approximate posterior from the VB solution, it can be seen that the two solutions have very similar forms. The main difference between the two solutions is that where VB has parameter means in the approximate posterior, the MCMC solution has the previous sample point in its full conditional distributions. The initial conditions for the MCMC solution are similar to the conditions for the VB solution, the mean parameters μ_J and μ_B are set to 1, and the precision parameters σ, λ and α are set to 1. The burn-in period was set to one million samples, to allow the chain to converge to the equilibrium distribution. After the burn-in period, one million more samples were generated to be used to estimate the parameters, a subset of these samples as well as the corresponding prior and likelihood values can be seen in Figure C.4. The Markov chain appears to have converged to the target distribution, with the chain exploring the entire sample space.

C.3.4 Results

Point estimates of the primary current density for a single time step can be seen in Figure C.5, in this plot weak sources (magnitude less than 25% of maximum) have been set to zero to emphasize the actual sources. The main difference between the two solutions is that the MCMC explores the full distribution, while the VB method hones in on a specific peak. As a result, the VB methods identifies a single source while the MCMC shows two possible sources. More specifically, the MCMC distribution suggests that there could be the activity could be due to the same positive source identified by VB, or it could also be a strong negative source from the opposite side of the brain.

The major difference between the VB and MCMC solution is in the estimates of parameter variance. When compared to the MCMC solution, the VB solution appears to greatly underestimate the variance of parameters in situations where there is an actual correlation between variables (see the example in Section C.1). It appears that this is also the case here, the difference between the variance of the VB and MCMC solutions for the important parameters is visualized in Figure C.6. There are very few negative values, meaning that the variance of the MCMC solution is typically larger than the VB solution

and these differences get as large as 160.

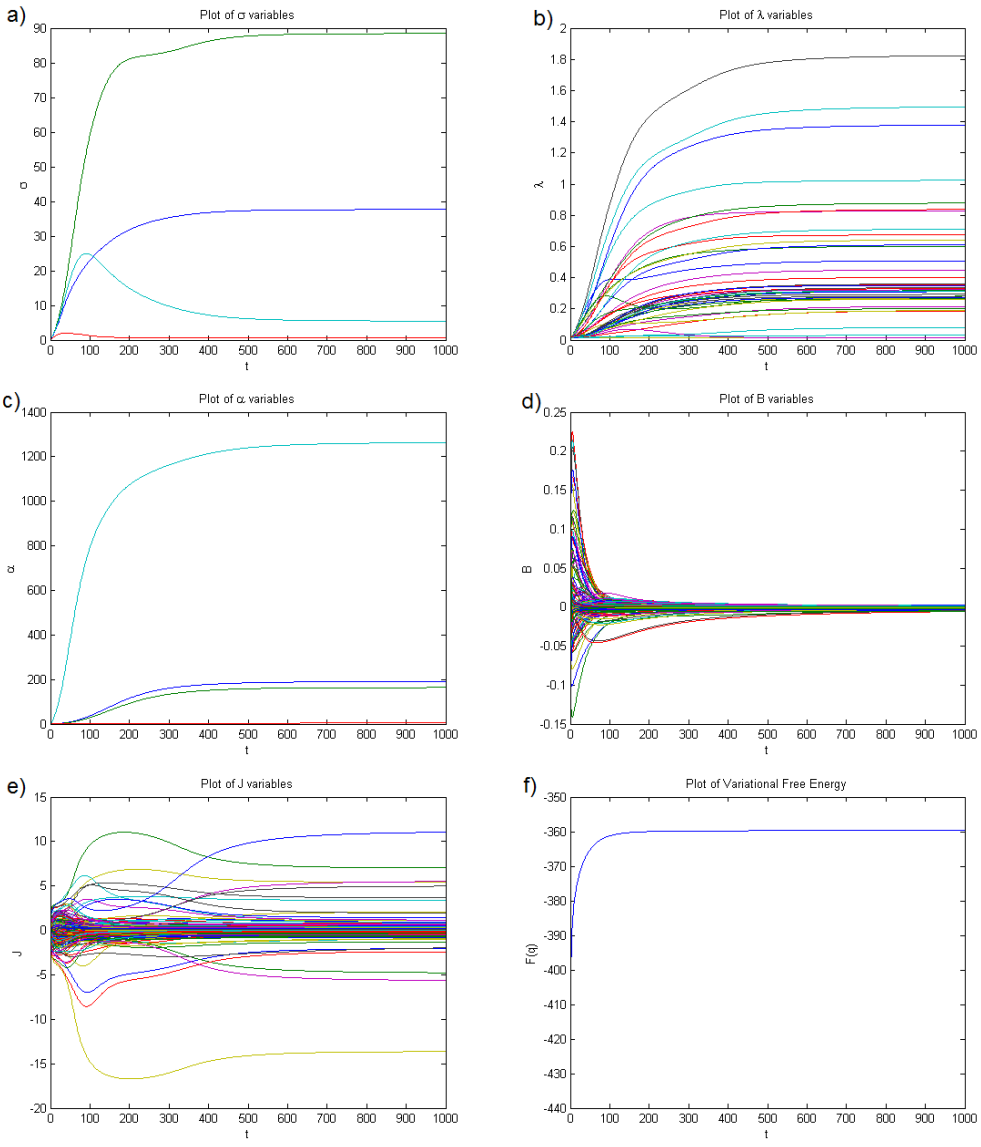

Figure C.3: Trajectory of the VB solution of the parameters a) σ, b) λ, c) α, d) B, e) J and f) F(q). The ARD prior lets the data determine the relevance of each regressor, generator and sensor. Figures a) and c) show that there is one specific regressor/sensor which is most important, while b) show a cluster of generators which are important. The plot plot f) shows that the VFE is converging to a maximum.

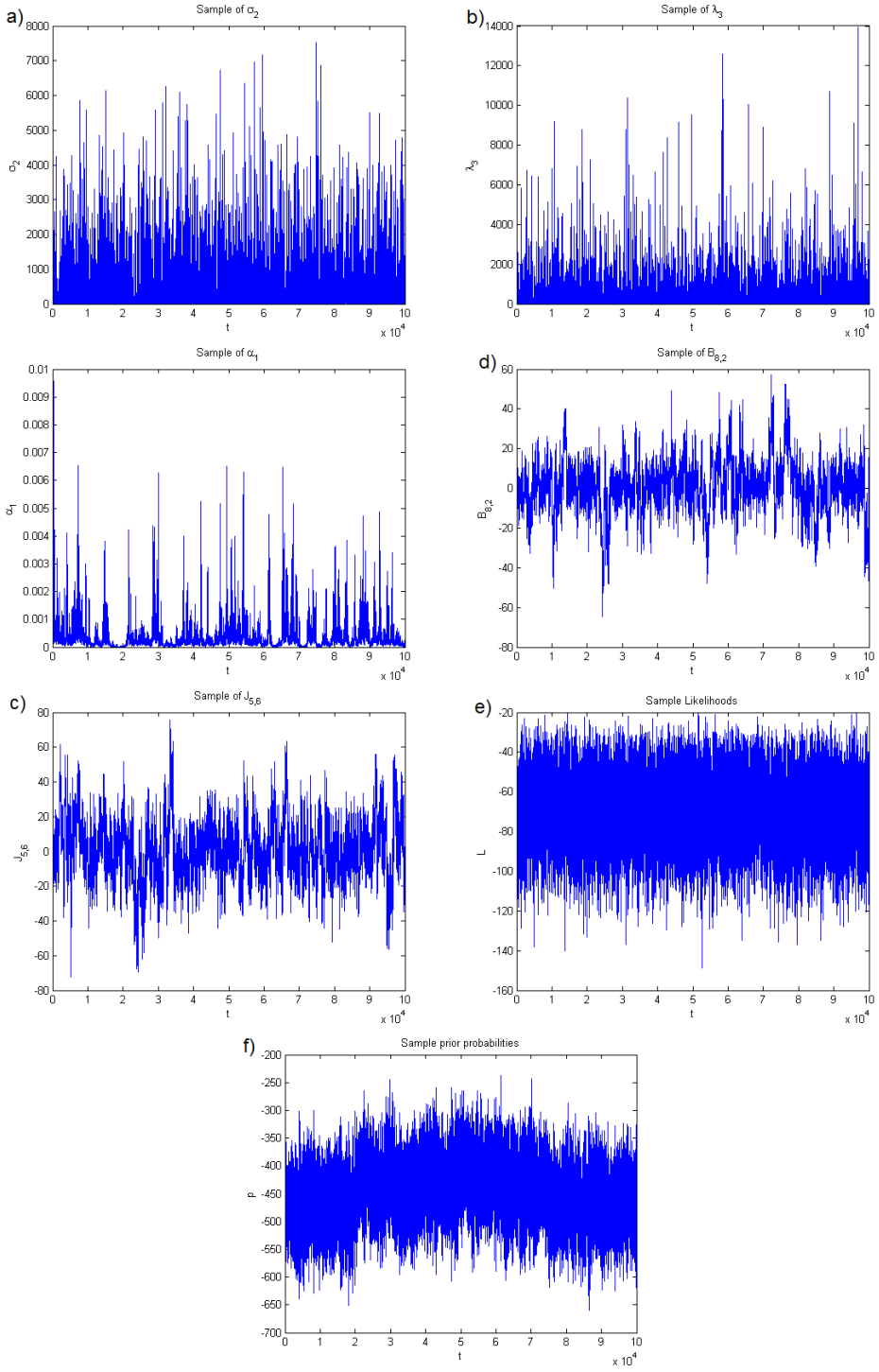

Figure C.4: Random samples from the posterior distribution generated using the MCMC solution, a) σ_2, b) λ_3, c) $J_{5,6}$ and d) $B_{6,2}$. Parts e) and f) show the sample likelihood and prior values respectively. The Markov chain appears to have converged to the target distribution, the chain is exploring the entire sample space.

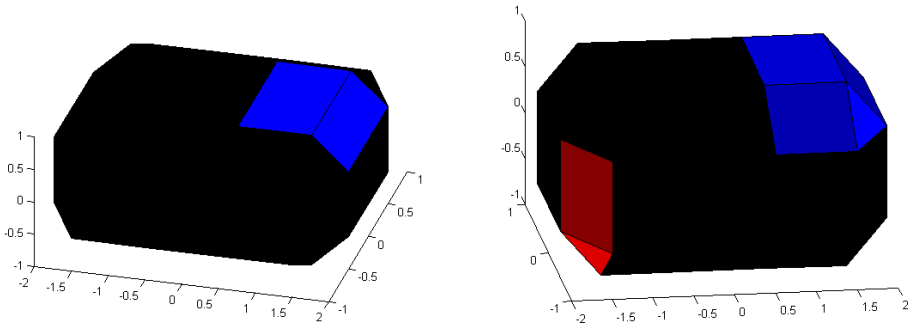

Figure C.5: Primary current density for the VB solution (left) and MCMC solution (right), currents with magnitude less than 25% of the maximum current density set to zero. The VB solution shows a single strong positive source at the location of the actual source density. The MCMC solution shows two possible current, one positive source at the location of the actual source and a negative source on the opposite side of the brain.

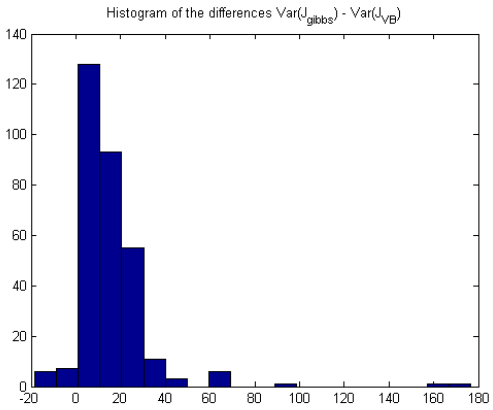

Figure C.6: A histogram of the difference between current density variance of the MCMC and VB solutions. There are very few negative values, which means that the variance of the MCMC solution is typically larger than the VB solution and the differences get as large as 160, meaning sometimes the VB solution greatly underestimates the variance of the posterior distribution.

Appendix D

Phase Shift Simulations

D.1 Proof of Theorem

A restatement of Theorem 3.1 and a proof of the results.

Theorem (3.1)**.** *Suppose x_t is a noisy oscillator sampled at T Hz with frequency f_0 Hz ($\omega = 2\pi f_0/T$) and constant phase ϕ i.e., $x_t = \sin(\omega t + \phi) + \epsilon_t$, with $E(\epsilon_t) = 0$ and $V(\epsilon_t) = \sigma^2$. Let $y_t = \mathbb{H}[x_t \sin\{-\omega t\}]$ and $\tilde{y}_t = \mathbb{H}[x_t \cos\{-\omega t\}]$ be the two components in the complex demodulation estimate of $\hat{\phi}_t$ (Eqn 1.2) with a EWMA(α) filter $\mathbb{H}[\cdot]$ given by (3.1). (a) The expected values of y_t and $\tilde{y}_t$ are*

$$E(y_t) = \frac{\cos(\phi)}{2} + b(y_t), \quad E(\tilde{y}_t) = \frac{\sin(\phi)}{2} + b(\tilde{y}_t)$$

where $b(y_t)$ and $b(\tilde{y}_t)$ are the biases of y_t and $\tilde{y}_t$, respectively. We have that

$$\begin{aligned} b(y_t) &= \frac{(1-\alpha)\big(\cos(2\omega t + \phi) + \alpha\cos(2\omega(t+1) + \phi)\big)}{2(1 - 2\alpha\cos(2\omega) + \alpha^2)} - \\ &\quad \frac{\alpha^{t+1}}{2}\left(\cos(\phi) - \frac{(1-\alpha)(\cos(\phi - 2\omega) - \alpha\cos(\phi))}{(1 - 2\alpha\cos(2\omega) + \alpha^2)}\right) \\ b(\tilde{y}_t) &= \frac{(1-\alpha)\big(\sin(2\omega t + \phi) - \alpha\sin(2\omega(t+1) + \phi)\big)}{2(1 - 2\alpha\cos(2\omega) + \alpha^2)} - \\ &\quad \frac{\alpha^{t+1}}{2}\left(\sin(\phi) + \frac{(1-\alpha)(\sin(\phi - 2\omega) - \alpha\sin(\phi))}{(1 - 2\alpha\cos(2\omega) + \alpha^2)}\right) \end{aligned}$$

(b) The variance of Y_t is

$$\begin{aligned} V(Y_t) &= \frac{\sigma^2}{2}\left(\frac{(1-\alpha)}{(1+\alpha)} - (1-\alpha)^2\right)\left(\frac{\cos(2\omega t) - \alpha^2\cos(2\omega(t+1))}{1-2\alpha^2\cos(2\omega)+\alpha^4}\right) - \\ &\quad \alpha^{2t+2}\left(\frac{(1-\alpha)}{(1+\alpha)} - \frac{(1-\alpha)^2(\cos(2\omega)-\alpha^2)}{1-2\alpha^2\cos(2\omega)+\alpha^4}\right) \end{aligned}$$

and for $\tilde{Y}_t$

$$\begin{aligned} V(\tilde{Y}_t) &= \frac{\sigma^2}{2}\left(\frac{(1-\alpha)}{(1+\alpha)} + (1-\alpha)^2\left(\frac{\cos(2\omega t) - \alpha^2\cos(2\omega(t+1))}{1-2\alpha^2\cos(2\omega)+\alpha^4}\right) - \right. \\ &\quad \left. \alpha^{2t+2}\left(\frac{(1-\alpha)}{(1+\alpha)} + \frac{(1-\alpha)^2(\cos(2\omega)-\alpha^2)}{1-2\alpha^2\cos(2\omega)+\alpha^4}\right)\right) \end{aligned}$$

(c) The covariance between Y_t and $\tilde{Y}_t$ is given by

$$Cov(Y_t, \tilde{Y}_t) = \frac{\sigma^2(1-\alpha)^2}{2}\left(\frac{\sin(2\omega t) - \alpha^2\sin(2\omega(t+1)) + \alpha^{2t+2}\sin(2\omega)}{1-2\alpha^2\cos(2\omega)+\alpha^4}\right)$$

Proof. First expanding y_t,

$$\begin{aligned} y_t &= \mathbb{H}[x_t \sin(\omega t)] \\ &= (1-\alpha)\sum_{k=0}^{t}\alpha^k \sin(\omega(t-k)+\phi)\sin(\omega(t-k)) + \epsilon_{t-k}\sin(\omega(t-k)). \end{aligned}$$

The expected value of y_t,

$$
\begin{aligned}
E(y_t) &= E\left((1-\alpha)\sum_{k=0}^{t}\alpha^k \sin\left(\omega(t-k)+\phi\right)\sin\left(\omega(t-k)\right) + \epsilon_{t-k}\sin\left(\omega(t-k)\right)\right) \\
&= (1-\alpha)\sum_{k=0}^{t}\alpha^k \sin\left(\omega(t-k)+\phi\right)\sin\left(\omega(t-k)\right) \\
&= (1-\alpha)\sum_{k=0}^{t}\alpha^k \left(\frac{\cos\left(\phi\right)-\cos\left(2\omega+\phi\right)}{2}\right) \\
&= \frac{(1-\alpha)}{2}\left(\sum_{k=0}^{t}\alpha^k\cos\left(\phi\right) - \sum_{k=0}^{t}\alpha^k\cos\left(2\omega+\phi\right)\right) \\
&= \frac{\cos\left(\phi\right)(1-\alpha^{t+1})}{2} - \frac{(1-\alpha)}{2}\Big(\frac{\cos\left(2\omega t+\phi\right)-\alpha\cos\left(2\omega(t+1)+\phi\right)}{1-2\alpha\cos\left(2\omega\right)+\alpha^2}\Big) \\
&\quad + \frac{\alpha^{t+1}}{2}\Big((1-\alpha)\frac{\cos\left(\phi-2\omega\right)-\alpha\cos\left(\phi\right)}{1-2\alpha\cos\left(2\omega\right)+\alpha^2}\Big) \\
&= \frac{\cos\left(\phi\right)}{2} - \frac{(1-\alpha)}{2}\Big(\frac{\cos\left(2\omega t+\phi\right)-\alpha\cos\left(2\omega(t+1)+\phi\right)}{1-2\alpha\cos\left(2\omega\right)+\alpha^2}\Big) \\
&\quad - \frac{\alpha^{t+1}}{2}\Big(\cos\left(\phi\right) - (1-\alpha)\frac{\cos\left(\phi-2\omega\right)-\alpha\cos\left(\phi\right)}{1-2\alpha\cos\left(2\omega\right)+\alpha^2}\Big) \\
&= \frac{\cos\phi}{2} + b(y_t)
\end{aligned}
$$

And for the variance of y_t,

$$\begin{aligned}
Var(y_t) &= Var\left((1-\alpha)\sum_{k=0}^{t}\alpha^k \sin\left(\omega(t-k)+\phi\right)\sin\left(\omega(t-k)\right)+\epsilon_{t-k}\sin\left(\omega(t-k)\right)\right) \\
&= (1-\alpha)^2\sum_{k=0}^{t}\alpha^{2k}\sin^2\left(\omega(t-k)\right)Var(\epsilon_{t-k}) \\
&= \sigma^2(1-\alpha)^2\sum_{k=0}^{t}\alpha^{2k}\sin^2\left(\omega(t-k)\right) \\
&= \sigma^2(1-\alpha)^2\sum_{k=0}^{t}\alpha^{2k}\frac{1-\cos\left(2\omega(t-k)\right)}{2} \\
&= \frac{\sigma^2(1-\alpha)^2}{2}\Big(\sum_{k=0}^{t}\alpha^{2k}-\sum_{k=0}^{t}\cos\left(2\omega(t-k)\right)\Big) \\
&= \frac{\sigma^2}{2}\left(\frac{(1-\alpha)(1-\alpha^{2t+2})}{(1+\alpha)}-(1-\alpha)^2\frac{(\cos\left(2\omega\right)-\alpha^2\cos\left(2\omega(t+1)\right)-\alpha^{2t+2}\cos\left(2\omega\right)+\alpha^{2t+4})}{(1-2\alpha^2\cos\left(2\omega\right)+\alpha^4)}\right) \\
&= \frac{\sigma}{2}\Bigg(\frac{1-\alpha}{1+\alpha}-(1-\alpha)^2\left(\frac{\cos\left(2\omega\right)-\alpha^2\cos\left(2\omega(t+1)\right)}{1-2\alpha^2\cos\left(2\omega\right)+\alpha^4}\right) \\
&\quad -\alpha^{2t+2}\left(\frac{1-\alpha}{1+\alpha}-(1-\alpha)^2\frac{\cos\left(-2\omega\right)-\alpha^2}{1-2\alpha^2\cos\left(2\omega\right)+\alpha^4}\right)\Bigg)
\end{aligned}$$

Next expanding $\tilde{y}_t$,

$$\begin{aligned}
\tilde{y}_t &= \mathbb{H}[x_t\cos\left(\omega t\right)] \\
&= (1-\alpha)\sum_{k=0}^{t}\alpha^k\sin\left(\omega(t-k)+\phi\right)\cos\left(\omega(t-k)\right)+\epsilon_{t-k}\cos\left(\omega(t-k)\right).
\end{aligned}$$

The expected value of $\tilde{y}_t$,

$$
\begin{aligned}
E(\tilde{y}_t) &= E\left((1-\alpha)\sum_{k=0}^{t}\alpha^k \sin\left(\omega(t-k)+\phi\right)\cos\left(\omega(t-k)\right) + \epsilon_{t-k}\cos\left(\omega(t-k)\right)\right) \\
&= (1-\alpha)\sum_{k=0}^{t}\alpha^k \sin\left(\omega(t-k)+\phi\right)\cos\left(\omega(t-k)\right) \\
&= (1-\alpha)\sum_{k=0}^{t}\alpha^k \left(\frac{\sin\left(\phi\right)+\sin\left(2\omega+\phi\right)}{2}\right) \\
&= \frac{(1-\alpha)}{2}\left(\sum_{k=0}^{t}\alpha^k \sin\left(\phi\right) - \sum_{k=0}^{t}\alpha^k \sin\left(2\omega+\phi\right)\right) \\
&= \frac{\sin\left(\phi\right)(1-\alpha^{t+1})}{2} - \frac{(1-\alpha)}{2}\Big(\frac{\sin\left(2\omega t+\phi\right)-\alpha\sin\left(2\omega(t+1)+\phi\right)}{1-2\alpha\sin\left(2\omega\right)+\alpha^2}\Big) \\
&\quad + \frac{\alpha^{t+1}}{2}\Big((1-\alpha)\frac{\sin\left(\phi-2\omega\right)-\alpha\sin\left(\phi\right)}{1-2\alpha\sin\left(2\omega\right)+\alpha^2}\Big) \\
&= \frac{\sin\left(\phi\right)}{2} + \frac{(1-\alpha)}{2}\Big(\frac{\sin\left(2\omega t+\phi\right)-\alpha\sin\left(2\omega(t+1)+\phi\right)}{1-2\alpha\cos\left(2\omega\right)+\alpha^2}\Big) \\
&\quad - \frac{\alpha^{t+1}}{2}\Big(\sin\left(\phi\right) + (1-\alpha)\frac{\sin\left(\phi-2\omega\right)-\alpha\sin\left(\phi\right)}{1-2\alpha\cos\left(2\omega\right)+\alpha^2}\Big) \\
&= \frac{\sin\phi}{2} + b(\tilde{y}_t)
\end{aligned}
$$

And for the variance of $\tilde{y}_t$,

$$\begin{aligned}
Var(\tilde{y}_t) &= Var\left((1-\alpha)\sum_{k=0}^{t}\alpha^k \sin\left(\omega(t-k)+\phi\right)\cos\left(\omega(t-k)\right) + \epsilon_{t-k}\cos\left(\omega(t-k)\right)\right) \\
&= (1-\alpha)^2\sum_{k=0}^{t}\alpha^{2k}\cos^2\left(\omega(t-k)\right)Var(\epsilon_{t-k}) \\
&= \sigma^2(1-\alpha)^2\sum_{k=0}^{t}\alpha^{2k}\cos^2\left(\omega(t-k)\right) \\
&= \sigma^2(1-\alpha)^2\sum_{k=0}^{t}\alpha^{2k}\frac{1+\cos\left(2\omega(t-k)\right)}{2} \\
&= \frac{\sigma^2(1-\alpha)^2}{2}\Big(\sum_{k=0}^{t}\alpha^{2k} + \sum_{k=0}^{t}\cos\left(2\omega(t-k)\right)\Big) \\
&= \frac{\sigma^2}{2}\left(\frac{(1-\alpha)(1-\alpha^{2t+2})}{(1+\alpha)} + (1-\alpha)^2\frac{(\cos(2\omega) - \alpha^2\cos\left(2\omega(t+1)\right) - \alpha^{2t+2}\cos(2\omega) + \alpha^{2t+4})}{(1-2\alpha^2\cos(2\omega)+\alpha^4)}\right) \\
&= \frac{\sigma}{2}\Bigg(\frac{1-\alpha}{1+\alpha} + (1-\alpha)^2\left(\frac{\cos(2\omega) - \alpha^2\cos\left(2\omega(t+1)\right)}{1-2\alpha^2\cos(2\omega)+\alpha^4}\right) \\
&\quad -\alpha^{2t+2}\left(\frac{1-\alpha}{1+\alpha} + (1-\alpha)^2\frac{\cos(-2\omega)-\alpha^2}{1-2\alpha^2\cos(2\omega)+\alpha^4}\right)\Bigg)
\end{aligned}$$

Finally the covariance between y_t and $\tilde{y}_t$,

$$\begin{aligned}
Cov(y_t, \tilde{y}_t) &= Cov\left((1-\alpha)\sum_{k=0}^{t}\alpha^k x_{t-k}\sin\left(\omega(t-k)\right), (1-\alpha)\sum_{k=0}^{t}\alpha^k x_{t-k}\cos\left(\omega(t-k)\right)\right) \\
&= Cov\left((1-\alpha)\sum_{k=0}^{t}\alpha^k e_{t-k}\sin\left(\omega(t-k)\right), (1-\alpha)\sum_{k=0}^{t}\alpha^k \epsilon_{t-k}\cos\left(\omega(t-k)\right)\right) \\
&= (1-\alpha)^2\sum_{k=0}^{t}\alpha^{2k}\sin\left(\omega(t-k)\right)\cos\left(\omega(t-k)\right)Var(\epsilon_{t-k}) \\
&= \sigma^2(1-\alpha)^2\sum_{k=0}^{t}\alpha^{2k}\sin\left(\omega(t-k)\right)\cos\left(\omega(t-k)\right) \\
&= \sigma^2(1-\alpha)^2\left(\frac{\sin\left(2\omega t\right) - \alpha^2\sin\left(2\omega(t+1)\right) + \alpha^{2t+2}\sin\left(2\omega\right)}{1 - 2\alpha^2\cos\left(2\omega\right) + \alpha^4}\right)
\end{aligned}$$

□

D.2 Estimator Properties

To illustrate the properties of the CUSUM and PD method using a 4^{th} order Butterworth filter, bootstrapped datasets were generated with a SNR of 0 dB (r=0.5), a length of 5 seconds ($N = 5T + N_{burn}$) and initial phase of zero ($\phi_0 = 0$); histograms of sample statistics, as well as qq-plots comparing the samples to fitted GEV distributions are shown in Figure D.1, the GEV distribution fits well on both statistics. Importantly, the GEV distributions are found to be independent of the specific value of ϕ_0.

Starting with an conservatively large value of N_{burn}, the sampling distribution can be established and the critical value Φ_α calculated. The value of N_{burn} can then be calibrated to be the minimum value such that

$$P(S_{1/2} > \Phi_\alpha) \leq \alpha.$$

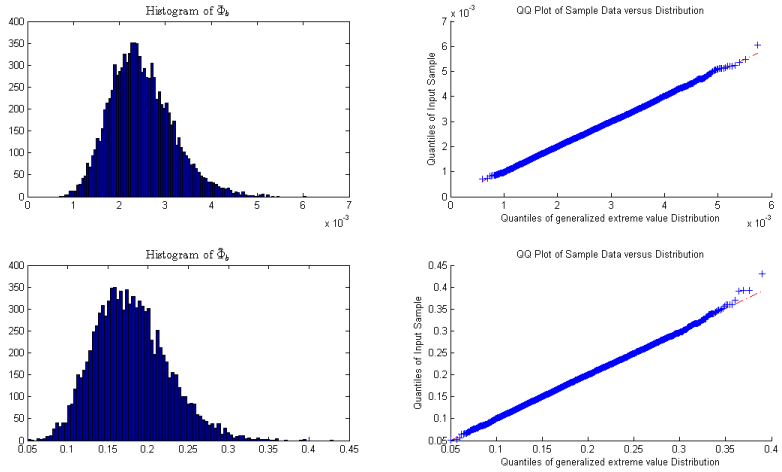

Figure D.1: Histograms (left) and qq-plots (right) for S_1 (top) and S_2 (bottom). Both histograms are positively skewed, though at quite different scales. The qq-plots suggest that the statistics are well describe by the GEV distributions.

D.3 Power Analysis

The parametric bootstrapping techniques can also be used to learn about the minimum shift magnitude Δ_{min} such that the shift events can be identified over the noise in the signal. To do this, bootstrapped datasets are simulated as above, but rather than a constant phase we include a single phase shift event at time t_1 with magnitude Δ,

$$\phi_t = \phi_0 + \Delta H(t - t_1).$$

The power of each test to identify a phase shift event is estimated for a range of SNR and Δ parameters, the results of this analysis are shown in Figure D.2. For low levels of noise, both methods are able to accurately identify phase shift events with a magnitude as small as $\Delta_{min} = 0.15$ radians. As the SNR decreases, the value of Δ_{min} increases, faster in S_2 than in S_1.

To quantify the temporal resolution of the PD estimator, we use bootstrapped datasets

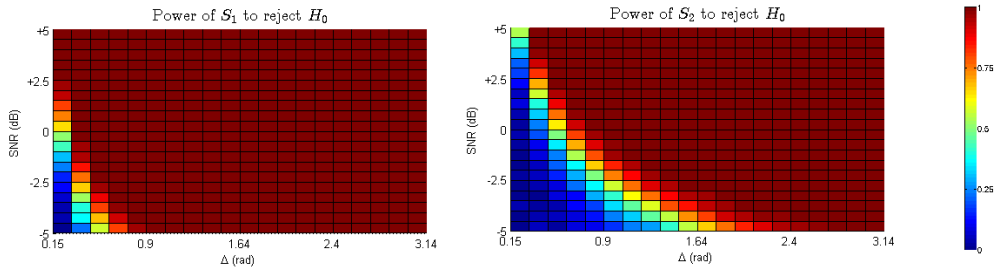

Figure D.2: Power to identify shift events of magnitude $\Delta\phi$ as a function of SNR. Both methods are able to identify shift events with magnitude as low as 0.15 radians. As the SNR increases, the range of identifiable shifts decreases in both statistics, but faster in S_2.

with two phase shift events,

$$\phi_t = \phi_0 + \Delta_1 H(t - t_0) + \Delta_2 H(t - t_0 - ISI).$$

We consider the power of each estimator to successfully identify both phase shift events as a function of the SNR and common shift magnitude $\Delta_1 = \Delta_2 = \Delta$ parameters.

For each estimator, the minimum ISI values are visualized in Figure D.3 for a range of SNR (r) and shift magnitude (Δ) parameters. Like the power analysis from the single shift simulations, the CUSUM estimator is able to reliably identify two phase shift events for a larger range of parameters than the PD estimator. For parameter values where shifts can be reliably identified for both estimators, the median ISI_{min} value of for the PD estimator is $884ms$, and for the CUSUM estimator it is $1024ms$. The PD estimator has the greater temporal resolution of the two methods, it is able to identify shifts $140ms$ closer together than the CUSUM estimator.

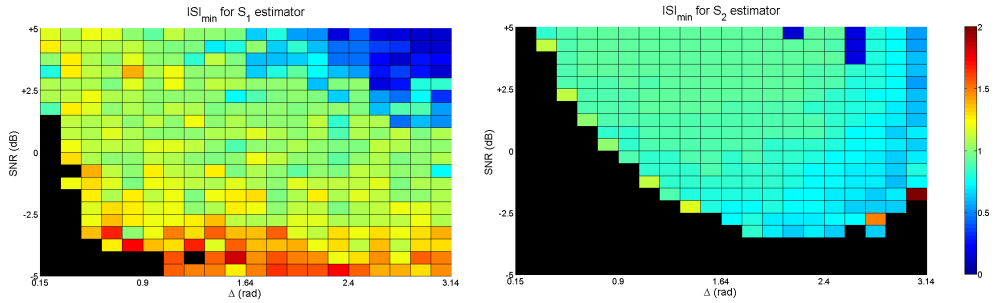

Figure D.3: Minimum ISI value for each estimator as function of SNR and shift magnitude. The CUSUM estimator is able to correctly identify two shifts for a larger range of parameters, but the PD estimator has the higher temporal resolution. For the range of parameters where both estimators can successfully identify two shifts, the median ISI_{min} value is $140ms$ shorter in the PD estimator.

Appendix E

Connectivity Measures

E.1 Granger Causality

In Granger's original paper, he introduced a statistical test for causality between two processes based on a bi-variate auto-regressive model [116]. The method has a convenient spectral representation which was later formalized as *directed coherence* (DC). Analysis using DC was not applied to neuroscience until over a decade later later [241].

An extension of Granger's method is the *directed transfer function* (DTF) which allows causality to be tested in a full multivariate system [154]. Another extension is *partial directed coherence* (PDC), which tests for causality while taking into account the effect of other possible variables [15]. The method of PDC has been applied to many neuroscience problems in recent years [268, 245, 11].

Granger's formation of causality assumes a stationary stochastic process, however there is evidence that this is not the case in EEG recordings [55, 76, 294]. This has led to the development of time dependent measures of Granger causality for use in neuroscience [9, 189, 140].

Directed Coherence

Definition E.1. *The Spectral density matrix of a stationary M variable time series is an M by M matrix S, where the $S_{i,j}$ entry is the CSD between $x^{(i)}$ and $x^{(j)}$*

$$S_{ij}(f) = \frac{X^{(i)}(f)\tilde{X}^{(j)}(f)}{N}$$

The DC is based on a factorization of S(f) [16],

$$S(f) = H(f)\Sigma H(f)^{\mathbf{H}}$$

where $H^{\mathbb{H}}$ is the Hermitian transpose of H. In practice, calculation of H(f) (and hence DC) makes use of a multivariate auto-regressive (MVAR) time series model.

Definition E.2. *If $\vec{x}_t$ follows a MVAR(p) model with diagonal covariance matrix Σ and coefficient matrices A_i, then the directed coherence from $x_t^{(i)}$ to $x_t^{(j)}$ is*

$$\gamma_{ij}(f) = \frac{H_{ij}(f)}{\sqrt{S_{ii}(f)}}$$

where

$$H(f) = (I - A(f))^{-1}$$

and

$$A(f) = \sum_{j=1}^{p} A_j e^{-i2\pi jf}$$

The PDC between two signals is based on a similar decomposition of the partial coherence function, where the influence of the other signals is accounted for [15].

Definition E.3. *If $\vec{x}_t$ follows a MVAR(p) model with diagonal Σ, then the partial directed coherence from $x_t^{(i)}$ to $x_t^{(j)}$ is*

$$\pi_{ij}(f) = \frac{\bar{A}_{ij}(f)}{\sqrt{\bar{a}_j^H(f)\Sigma^{-1}\bar{a}_j(f)}}$$

where $\bar{a}_j(f)$ are the columns of

$$\bar{A}(f) = I - \sum_{j=1}^{p} A_j e^{-i2\pi jf}$$

Consider the following example of a MVAR(2) model for N=3 signals,

$$x_t^{(1)} = 0.5x_{t-1}^{(1)} - 0.2x_{t-2}^{(1)} + 0.6x_{t-1}^{(3)} \quad \text{(E.1)}$$
$$x_t^{(2)} = 0.8x_{t-1}^{(1)} + 0.3x_{t-1}^{(2)} \quad \text{(E.2)}$$
$$x_t^{(3)} = -0.3x_{t-1}^{(3)} + 0.4x_{t-1}^{(1)} + 0.05x_{t-2}^{(1)} \quad \text{(E.3)}$$

The DC and PDC for this MVAR model are shown in Figure E.1, the $(i,j)^{th}$ plot represents $\gamma_{ij}(f)$ (or $\pi_{ij}(f)$). Both methods identify that $x_t^{(2)}$ does not contribute to any other signal, the DC and PDC is zero in the (1,2) and (3,2) plots. The difference between the two methods is most apparent when considering plot (2,3) the effect of x_3 on x_2, DC shows an effect while PDC identifies that x_3 only effects x_2 through x_1 and thus there is no effect.

E.2 Mutual Information

Entropy is a fundamental quantity used to describe the uncertainty associated with a random system. The fields of information theory [254], [253], [206] and thermodynamics [32], [108] both use the same form of entropy measure, often called the Shannon-Boltzmann-Gibbs (SBG) entropy.

Definition E.4. *For a stationary discrete random variable X with possible states $\{x_i\}_{i=1}^N$ and associated probability mass function $P_X(x_i)$ the SBG entropy is defined as*

$$H(X) = -\sum_{i=1}^{N} P_X(x_i)\log_2(P_X(x_i)).$$

Definition E.5. *For a stationary continuous random variable X with support S and probability density function f(x), the SBG entropy is defined as*

$$H(X) = -\int_S f(x)\log_2(f(x))dx.$$

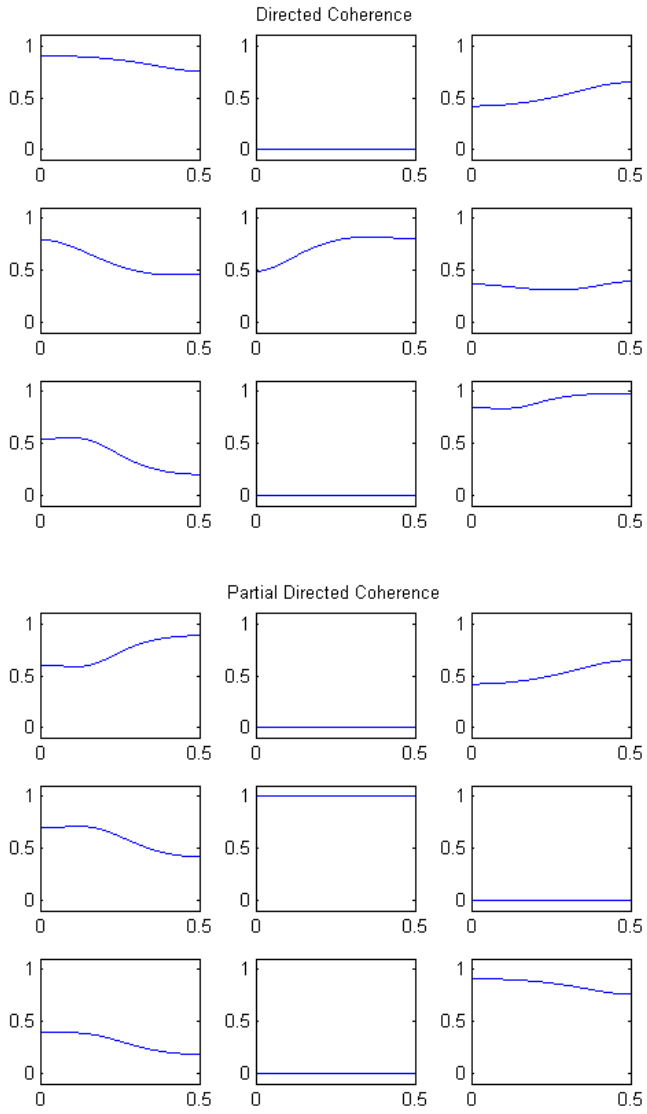

Figure E.1: Analysis of the above model using DTF (top) and PDC (bottom) forms of Granger causality.

The units of a measured entropy is determined by the base of the logarithm, when the base is two then the units are bits.

When there is more than one signal, the idea of entropy naturally generalizes to the multivariate case.

Definition E.6. *The joint entropy of two discrete stationary processes X_t and Y_t is*

$$H(X,Y) = -\sum_{i,j=1}^{N,M} P_{X,Y}(x_i, y_j) \log_2(P_{X,Y}(x_i, y_j)).$$

Entropy can be used to describe the statistical dependence between two signals using *mutual information* (MI). When two signals have a statistical dependence, observing one signal reveals information about the other due to their common structure. The MI between two signals is the reduction in entropy caused by knowledge of the other.

Definition E.7. *The MI is between two signals X and Y with possible states $\{x_i\}_1^N$ and $\{y_j\}_1^M$ is defined as,*

$$\begin{aligned} MI(X,Y) &= \sum_{i,j=1}^{N,M} P_{X,Y}(x_i, y_j) \log \frac{P_{X,Y}(x_i, y_j)}{P_X(x_i) P_Y(y_j)} && \text{(E.4)} \\ &= H(X) + H(Y) - H(X,Y) && \text{(E.5)} \end{aligned}$$

The MI between two processes can also be thought of in terms of *Kullback-Leibler* (KL) divergence.

Definition E.8. *If p and q are two probability distributions and $q = 0$ implies that $p = 0$, then the KL divergence from p to q is defined as*

$$KL(p||q) = \int_x p(x) \log \frac{p(x)}{q(x)} dx.$$

The KL divergence is a non-negative quantity, and $KL(p||q) = 0 \; iff \; p \equiv q$. The MI between random variables X and Y can be though of as the KL divergence of the product of marginal distribution from the joint probability distribution,

$$MI(X,Y) = KL(f(x,y)||f_x(x) f_y(y))$$

Thus the MI is equal to zero if and only if the random variables are independent.

Mutual information analysis has been used recently to assess connectivity between signals in EEG recordings. MI has the ability to detect any type of statistical dependence, where as coherence is a measure based on linear dependence. This method was first applied to EEG recordings by Xu in 1997 [300]. Since then it has been applied to EEG studies of Alzheimer's [148], schizophrenia [196], sleep deprivation [195], aging [228] and Parkinson's [170], [293].

Information transmission across the brain is not instantaneous, so EEG analysis focuses on lagged MI quantities of the form

$$MI(X_t, Y_{t+\tau}), \quad 0 \leq \tau,$$

where τ is some non-negative integer. For values of $\tau \geq 1$ MI provides an asymmetric measure of information flow.

Appendix F

PSGC Results

F.1 Individual Connectivity Maps

The primary physiological result of the PSGC method on the real EEG data were task differences in long range connectivity and clustering coefficient measures on individual networks. Here we include plots of the connectivity maps for each of the 18 participants in the EEG study. The maps are not exactly to scale, the eight circles on each hemisphere represent the eight electrodes used in the study (e.g., for the left hemisphere Fp1, F7, T3, T5, O1 laterally; F3, C3, P3 medially). The nose is pointed to the top of the circles. Resting networks are shown in Figure F.1, auditory vigilance in Figure F.2 and visual vigilance in Figure F.3.

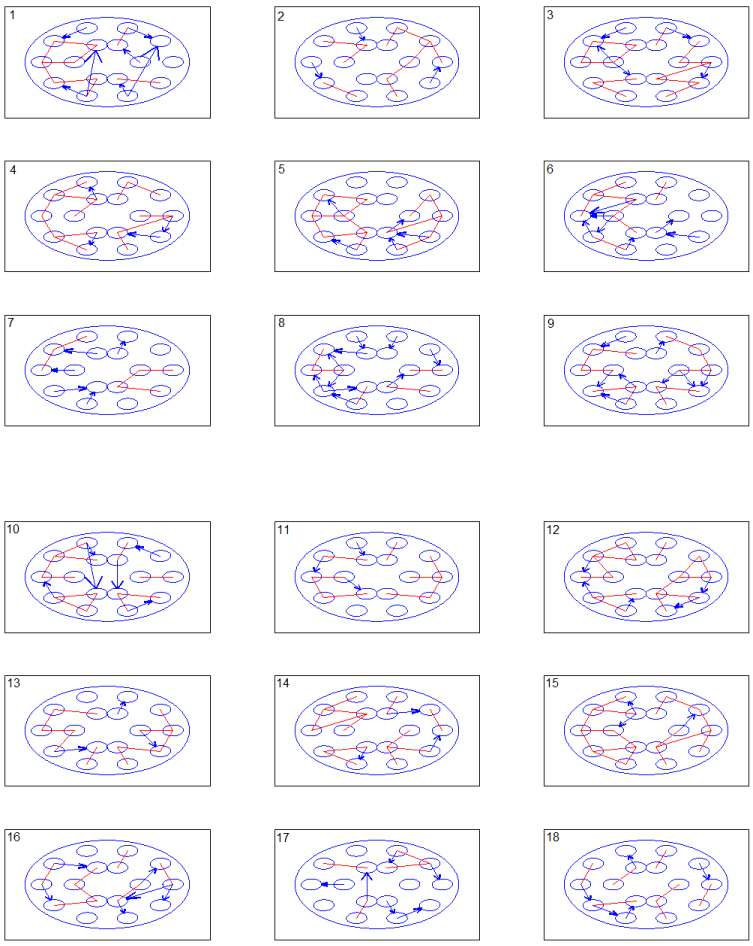

Figure F.1: Estimated individual connectivity networks for each participant in the resting task. The connections in the resting task are predominantly short range, connecting neighbouring locations.

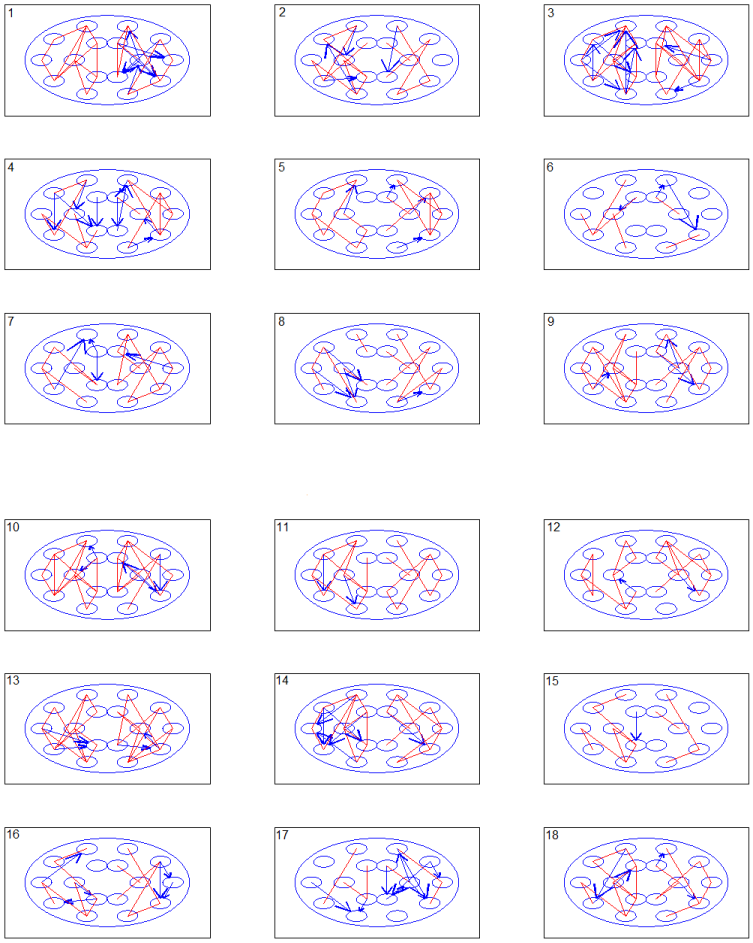

Figure F.2: Estimated individual connectivity networks for each participant in the auditory vigilance task. In both the auditory and visual vigilance tasks we see higher rates of long range connectivity, specifically connecting the prefrontal and parietal locations.

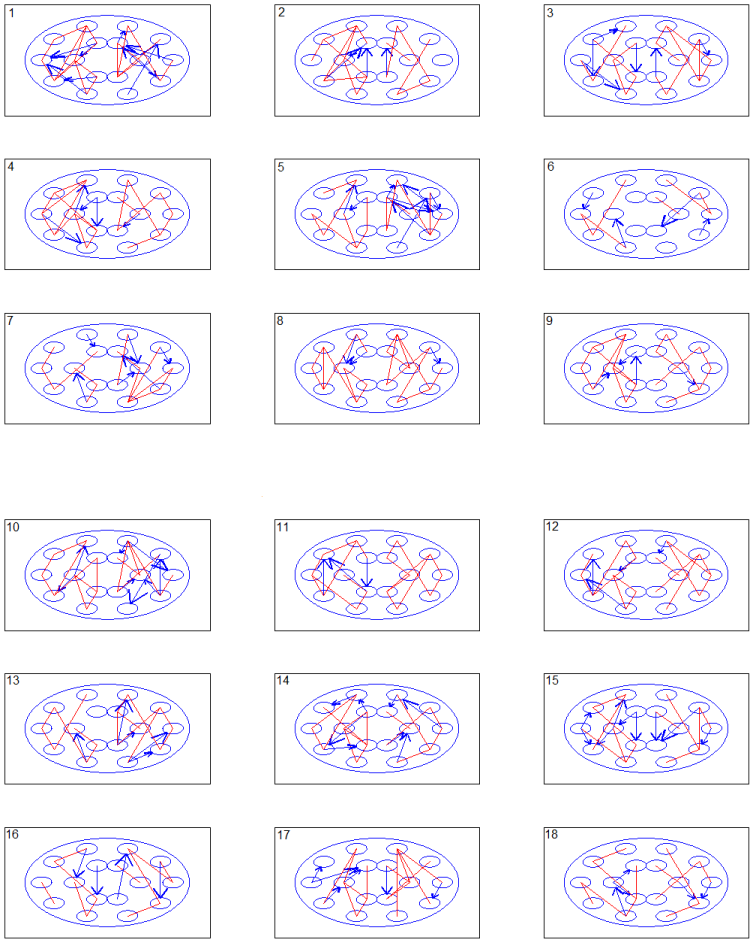

Figure F.3: Estimated individual connectivity networks for each participant in the visual vigilance task. In both the auditory and visual vigilance tasks we see higher rates of long range connectivity, specifically connecting the prefrontal and parietal locations.

Appendix G

Complex Demodulation

The complex demodulation algorithm estimates the instantaneous phase of a time series (x_t) at a specific frequency f_0 [28, 113]. First two new time series are defined which represent the original and Hilbert transformed signals (the cosine is a $\pi/2$ phase shift of sine), with frequency components shifted to 0Hz,

$$y_t = x_t \sin\left(\frac{2\pi f_0 t}{T}\right)$$

$$\tilde{y_t} = x_t \cos\left(\frac{2\pi f_0 t}{T}\right).$$

Note that only one of the two components of the oscillator ($+f_0$, $-f_0$) is shifted towards zero, the other is shifted to $2f_0$. Additionally, only the component at zero is a true $\pi/2$ phase shift; the $2f_0$ component must be removed so that $\tilde{y}$ represents the Hilbert transform. Conditions for a signal to have a physically meaningful representation in terms of instantaneous phase and amplitude are that it be mono-component, (locally described by a frequency components contained in a small band) and *narrowband*, (relative changes in amplitude are small compared to changes in phase) [48],

$$\left|\frac{d\phi_t}{dt}\right| >> \left|\frac{1}{A_t}\frac{dA_t}{dt}\right|$$

To ensure these conditions are satisfied, the signal is put through a δHz low-pass filter; this removes the component at $2f_0$, as well as any other unwanted components (possibly

due to noise). What remains is power in a small band $(f_0-\delta, f_0+\delta)$ from the original signal.

The instantaneous phase and instantaneous magnitude are then calculated as,

$$A_t = \sqrt{(y_t)^2 + (\tilde{y}_t)^2}$$

$$\phi_t = \tan^{-1}\left(\frac{y_t}{\tilde{y}_t}\right)$$

An example of the complex demodulation algorithm, consider the following signal with $4Hz$ and $12Hz$ components, and a change-point at $t = 500$,

$$x_t = \begin{cases} 4\sin\left(\frac{2\pi 12t}{T} + 1\right) + 2\sin\left(\frac{2\pi 4(t-14)}{T}\right) & \textit{if} \quad t \leq 500 \\ 8\sin\left(\frac{2\pi 12t}{T} + 2\right) + 2\sin\left(\frac{2\pi 4(t-14)}{T}\right) & \textit{otherwise.} \end{cases}$$

In this example the component of interest is the $12Hz$ oscillator; the signal is put through a $\delta = 2Hz$ low-pass filter to remove components outside the ($10Hz$-$14Hz$) band. In Figure G.1 are plots of the actual and estimated instantaneous phase and amplitude of the 12Hz component, in this idealized scenario the method produces accurate estimate with only slight boundary effects due to filtering, at the beginning and around the change point in the signal. If the signal is not run through the filter then it will not satisfy the conditions necessary for an interpretable result. Figure G.1 also shows the results of calculating the instantaneous phase from the original signal without first filtering out the contributions from other frequency bands. The results no longer correspond to the instantaneous phase and amplitude of the 12Hz component.

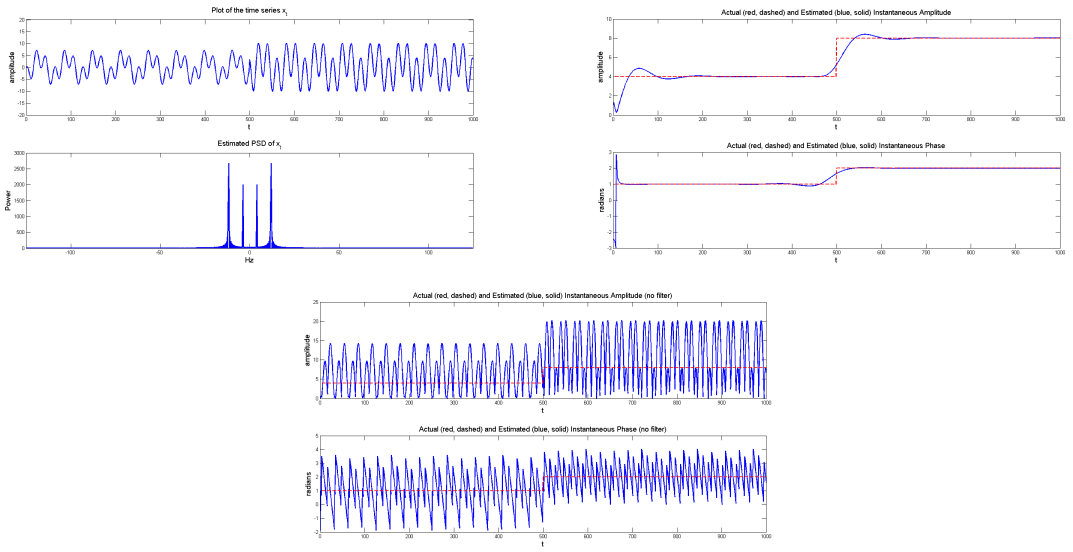

Figure G.1: Top: Estimated instantaneous amplitude of the original signal (blue) and actual amplitude (red). Bottom: Estimated instantaneous phase of the original signal (blue) and actual phase (red). In the unfiltered situation, the 4Hz component shows up explicitly in the estimated instantaneous amplitude and phase.

Nomenclature

(f)MRI (functional) magnetic resonance imaging

ACF auto-correlation function

AR auto regressive

ARD automatic relevance detection

AUROC area under the ROC

BCI brain-computer interface

bSNR biological signal-to-noise ratio

CIF conditional intensity function

CPA change point analysis

CSD cross-spectral density

CUSUM cumulative summation

DC directed coherence

DC distributed current

DCM dynamic causal modelling

DTF directed transfer function

EEG electroencephalogram

EM expectation maximization

ERP event related potential

EWMA exponentially weighted moving average

FFT fast Fourier transform

FIR finite impulse response

FN false negative

FP false positive

GC Granger causality

GEV generalized extreme value

ICA independent component anlaysis

IIR infinite impulse response

ISI inter-shift interval

ITI inter-trial interval

KL kullbeck-leibler

LARS least angel regressor selection

LASSO local auto shrinkage and selection operator

LFM lead field matrix

LRT likelihood ratio test

mACC maximum accuracy

MAP maximum *a posteriori*

MCMC markov chain monte carlo

MI mutual information

MNE minimum norm estimate

MVAR multivariate auto-regressive

PD phase derivative

PDC partial directed coherence

PDD phase difference derivative

PLV phase locking value

PR phase reset

PSD power spectral density

PSGC phase shift Granger causality

PSI phase slope index

ROC receiver operator curve

ROI region of interest

sLORETA standardized low resolution tomography

SNR signal-to-noise ratio

SOC self organized criticality

SR source reconstruction

TN true negative

TP true positive

VB variational Bayes

VC volume conduction

VEP visual evoked potential

VFE variational free energy

References

[1] U.R. Abeyratne, Y. Kinouchi, H. Oki, J. Okada, F. Shichijo, and K. Matsumoto. Artificial neural networks for source localization in the human brain. *Brain Topography*, 4:3–21, 1991.

[2] W.R. Adey and D.O. Walter. Application of phase detection and averaging techniques in computer analysis of EEG records in the cat. *Exp Neurol*, 7:186–209, 1963.

[3] E.D. Adrian and K. Yamagiwa. The origin of the berger rhythm. *Brain*, 58:323–351, 1935.

[4] H. Akaike. A new look at the statistical model identification. *IEEE Trans Autom Control*, 19:716–723, 1974.

[5] P. Andersen, J.C. Eccles, and Y. Løyning. Hippocampus of the brain: Recurrent inhibition in the hippocampus with identification of the inhibitory cell and its synapses. *Nature*, 198:540–542, 1963.

[6] J. Antoch and M. Hušková. Permutation tests in change point analysis. *Statistics and Probability Letters*, 53:37–46, 2001.

[7] J.P. Ary, S.A. Klein, and D.H. Fender. Location of sources of evoked scalp potentials: Corrections for skull and scalp thickness. *IEEE Trans Biomed Eng*, 28:447–452, 1981.

[8] L. Astolfi, F. Cincotti, D. Mattia, C. Babiloni, F. Carducci, A. Basilisco, P. M. Rossini, S. Salinari, L. Ding, Y. Ni, B. He, and F. Babiloni. Assessing cortical functional connectivity by linear inverse estimation and directed transfer function:

simulations and application to real data. *Clinical Neurophysiology*, 116:920–932, 2005.

[9] L. Astolfi, F. Cincotti, D. Mattia, F. De Vico Fallani, A. Tocci, A. Colosimo, S. Salinari, M.G. Marciani, W. Hesse, H. Witte, M. Ursino, M. Zavaglia, and F. Babiloni. Tracking the time-varying cortical connectivity patterns by adaptive multivariate estimators. *IEEE Trans Biomed Eng*, 55:902–913, 2008.

[10] L. Astolfi, F. Cincotti, D. Mattia, M. Grazia Marciani, L. A. Baccala, F. de Vico Fallani, S. Salinari, M. Ursino, M. Zavaglia, L Ding, J. C. Edgar, G. A. Miller, B. He, and F. Babiloni. Comparison of different cortical connectivity estimators for high-resolution EEG recordings. *Human Brain Mapping*, 28:143–157, 2007.

[11] L. Astolfi, J. Toppi, F. De Vico Fallani, G. Vecchiato, S. Salinari, D. Mattia, F. Cincotti, and F. Babiloni. Neuroelectrical hyperscanning measures simultaneous brain activity in humans. *Brain Topogr*, 23:243–356, 2010.

[12] H. Attias. Inferring parameters and structure of latent variable models by variational bayes. In *Proceedings of the Fifteenth Conference of Uncertainty in Artificial Intelligence*, 1999.

[13] T. Auranen, A. Nummenmaa, M.S. Hamalainen, I.P. Jaaskelainen, J. Lampinen, A. Vehtari, and M. Sams. Bayesian analysis of the neuromagnetic inverse problem with l-p norm priors. *NeuroImage*, 26:870–884, 2005.

[14] F. Babiloni, F. Cincotti, C. Babiloni, F. Carducci, D. Mattia, L. Astolfi, A. Basilisco, P. M. Rossini, L. Ding, Y. Ni, J. Cheng, K. Christine, J. Sweeney, and B. He. Estimation of the cortical functional connectivity with the multimodal integration of high-resolution EEG and fMRI data by directed transfer function. *NeuroImage*, 24:118–131, 2005.

[15] L.A. Baccala and K. Sameshima. Partial directed coherence: a new concept in neural structure determination. *Biol Cybern*, 84:463–474, 2001.

[16] L.A. Baccala, K. Sameshima, G. Ballester, A.C. Do Valle, and C. Timo-laria. Studying the interaction between brain structures via directed coherence and granger causality. *Applied Signal Processing*, 5:40–48, 1998.

[17] O. Bai, P. Lin, S. Vorbach, M. K. Floeter, N. Hattori, and M. Hallett. A high performance sensorimotor beta rhythm-based brain-computer interface associated with human natural motor behavior. *J. Neural Eng.*, 5:24–35, 2008.

[18] S. Baillet and L. Garnero. A bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem. *IEEE Transactions on Biomedical Engineering*, 44:374–385, 1997.

[19] P. Bak, C. Tang, and K. Wisenfeld. Self organized criticality - an explanation of 1/f noise. *Phys Rev Lett*, 59:381–384, 1987.

[20] R. Balocchi, D. Menicucci, E. Santarcangelo, L. Sebastiani, A. Gemignani, B. Ghelarducci, and M. Varanini. Deriving the respiratory sinus arrhythmia from the heartbeat time series using empirical mode decomposition. *Chaos, Solitons & Fractals*, 20:171–177, 2004.

[21] A-L. Barabasi and R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.

[22] R.C. Barr, T.C. Pilkington, J.P. Boineau, and M.S. Spach. Determining surface potentials from current dipoles with applications to electrocardiology. *IEEE Trans Biomed Eng*, 13:88–92, 1966.

[23] P. Berg and M. Scherg. A fast method for forward computation of multiple-shell spherical head models. *Electroencephalogr Clin Neurophysiol*, 90:58–64, 1994.

[24] A. Berger, O. Kofman, U. Livneh, and A. Henik. Multidisciplinary perspectives on attention and the development of self-regulation. *Prog Neurobiol*, 82:256–286, 2007.

[25] H. Berger. Uber das elektrenkephalogramm des menschen. *Archiv fur Psychiatrie und Nervenkrankheiten*, 87:527–570, 1929.

[26] G. K. Bhattacharyya and R. A. Johnson. Tests for shift at an unknown time point. *The Annals of Mathematical Statistics*, 39:1731–1743, 1968.

[27] R.G. Bickford. Electroencephalographic diagnosis of brain tumor. *The American Journal of Surgery*, 93:946–951, 1957.

[28] C. Bingham, M.D. Godfrey, and J.W. Tukey. Modern techniques of power spectrum estimation. *IEEE Trans on Audio and Electroacoustics*, AU-15:56–66, 1967.

[29] C.D. Binnie, P.A. Ward, and J. Heywood. EEG contour mapping I: Theory and practice. *Proc. Electrophysiol. Technol. Ass.*, 18, 1971.

[30] R. Biscay, M. Lavielle, A. González, I. Clark, and P. Valdés. Maximum a posteriori estimation of change points in the EEG. *International Journal of Bio-Medical Computing*, 38:189–196, 1995.

[31] S. Boccaletti, J. Kurths, G. Osipov, D.L. Valladares, and C.S. Zhou. The synchronization of chaotic systems. *Phys Rep*, 366:1–101, 2002.

[32] L. Boltzmann. Weitere studien uber das warmegleichgewicht unter gasmolekulen. *Wiener Berichte*, 66:275–370, 1872.

[33] J. Bosch-Bayard, P. Valdéz-Sosa, T. Virues-Alba, E. Aubert-Vázquez, E.R. John, T. Harmony, J. Riera-Díaz, and N. Trujillo-Barreto. 3D statistical mapping of EEG source spectraby means of variable resolution electromagnetic tomography (vareta). *Clin Electroencephalogr*, 32:47–61, 2001.

[34] V. Braitenberg. Cortical architectonics: General and areal. In M.A. Brazier and H. Petsche, editors, *Architectonics of the cerebral cortex*. Raven Press, 1978.

[35] V. Braitenberg. Charting the visual cortex. *Cerebral Cortex*, 3:379–414, 1985.

[36] M.A.B. Brazier. The electrical fields at the surface of the head during sleep. *Electroencephalogr Clin Neurophysiol*, 1:195–204, 1949.

[37] M. Breakspear, L.M. Williams, and C.J.Stam. A novel method for the topographic analysis of neural activity reveals formation and dissolution of 'dynamic cell assemblies'. *J Comp Neurol*, 16:49–68, 2004.

[38] S.L. Bressler, R. Coppola, and R. Nakamura. Episodic multiregion cortical coherence at multiple frequencies during visual task performance. *Nature*, 366:153–156, 1993.

[39] B. E. Brodsky and B. S. Darkhovsky. *Non-parametric statistical dianosis: problems and methods*. Kluwer Academic Publishers, 2000.

[40] B.E. Brodsky, B.S. Darkhovsky, A.Y. Kaplan, and S.L. Shishkin. A nonparametric method for the segmentation of the EEG. *Computer Methods and Programs in Biomedicine*, 60:93–106, 1999.

[41] D.A. Brody, F.H. Terry, and R.E. Ideker. Eccentric dipole in a spherical medium: generalized expression for surface potentials. *IEEE Trans Biomed Eng*, 141-143, 1973.

[42] T.J. Buschman and E.K. Miller. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science*, 315:1860–1862, 2007.

[43] G. Buzsaki and A. Draguhn. Neural oscillations in cortical networks. *Science*, 304:1926–1929, 2004.

[44] G. Byford. An index for magnetic tape i. *Med Biol Eng Comput*, 3:249–259, 1965.

[45] G.A. Calvert, P.C. Hansen, S.D. Iversen, and M.J. Brammer. Detection of audio-visual integration sites in humans by applications of electrophysiological criteria to the bold effect. *NeuroImage*, 14:427–438, 2001.

[46] L. Canuet, R. Ishii, R. D. Pascual-Marqui, M. Iwase, R. Kurimoto, Y. Aoki, S. Ikeda, H. Takahashi, T. Nakayuki, and M. Takeda. Resting-state EEG source localization and functional connectivity in schizophrenia-like psychosis of epilepsy. *PLoS One*, 6:e27863, 2011.

[47] G.E. Chatrian, L. Bergamini, M. Dondey, D.W. Klass, M. Lennox-Buchtal, and I. Petersen. Report of the committee on terminology. *Electroencephalogr Clin Neurophysiol*, 37, 1974.

[48] M. Chavez, M. Besserve, C. Adam, and J. Martinerie. Towards a proper estimation of phase sychronization from time series. *J Neurosci Methods*, 154:149–160, 2006.

[49] J. Chen and A. K. Gupta. *Parametric Statistical Change Point Analysis*. Birkhause, 2000.

[50] H. Chernoff and S. Zacks. Estimating the current mean of a normal distribution which is subject to changes in time. *The Annals of Mathematical Statistics*, 35:999–1018, 1964.

[51] R. Chialvo. Critical brain networks. *Physica A: Statistical Mechanics*, 340:756–765, 2004.

[52] L. Cimponeriu, M.G. Rosenblum, T. Fieseler, J. Dammers, M. Schiek, M. Majtanik, P. Morosan, A. Bezerianos, and P.A. Tass. Inferring asymmetric relations between interacting neural oscillators. *Prog Theor Phys Suppl*, 150:22–36, 2003.

[53] C.J.S. Clark. Probabilistic methods in a biomagnetic inverse problem. *Inverse Problems*, 5:999–1012, 1989b.

[54] C.J.S. Clark and B.S. Janday. The solution of the biomagnetic inverse problem by maximum statistical entropy. *Inverse Problems*, 5:483–500, 1989a.

[55] B.A. Cohen and A. Sances. Stationarity of the human electroencephalogram. *Med Biol Eng Comput*, 15:513–518, 1977.

[56] J.W. Cooley and J.W. Tukey. An algorithm for the machine calculation of complex fourier series. *Mathematics of Computation*, 19:297–301, 1965.

[57] R. Cooper, J.W. Osselton, and J.C. Shaw. *EEG Technology*. Butterworth, 1980.

[58] F. Lopes da Silva. Biological aspects of eeg and magnetoencephalogram generation. In E. Niedermeyer and F. Lopes da Silve, editors, *Electroencephalography : Basic Principles, Clinical Applications and Related Fields*, chapter 5. Urban and Schwarzenberg, 1987.

[59] D.J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes*. Springer, 2003.

[60] I. Daly, C. M. Sweeney-Reed, and S. J. Nasuto. Testing for significance of phase synchronisation dynamics in the EEG. *J Comput Neurosci*, 34:411–432, 2013.

[61] J. Daunizeau, C. Grova, G. Marrelec, J.Mattout, S. Jbabdi, M. Pelegrini-Isaac, J-M. Lina, and H. Benali. Symmetrical event-related EEG/fMRI information fusion in a variational bayesian framework. *2007*, 36:69–87, 2007.

[62] J. Daunizeau, J. Mattout, D. Clonda, B. Goulard, H. Benali, and J. Lina. Bayesian spatio-temporal approach for EEG source reconstruction: Conciliating ECD and distributed models. *IEEE Trans Biomed Eng*, 53:503–516, 2006.

[63] O. David, D. Cosmelli, and K. Friston. Evaluation of different measures of functional connectivity using a neural mass model. *NeuroImage*, 21:659–673, 2004.

[64] O. David and L. Garnero. Time-coherent expansion of MEG/EEG cortical source. *NeuroImage*, 17:1277–1289, 2002.

[65] H. Davis, P.A. Davis, A.L. Loomis, E.N. Harvey, and G. Hobart. Electrical reactions of the human brain to auditory stimulation during sleep. *J Neurophysiol*, 2:500–514, 1939.

[66] H. Davis, S.K. Hirsh, J. Shelnutt, and C. Bowers. Futher validation of evoked reponse audiometry. *J Speech Hearing Res*, 10:717–732, 1967.

[67] G.D. Dawson. Cerebral response to electrical stimulation of peripheral nerve in man. *J Neurol Neurosurg Psychiatry*, 10:134–140, 1947.

[68] G.D. Dawson and W.G. Walter. The scope and limitations of visual and automatic analysis of the electroencephalogram. *J Neurol Neurosurg Psychiatry*, 7:119–133, 1944.

[69] R. Grave de Peralta Menendez, O. Hauk, S. Gonzalez Ardino, H. Vogt, and C. Michel. Linear inverse solutions with optimal resolution kernels applied to the electromagnetic inverse. *Hum Brain Mapp*, 5:454–467, 1997.

[70] R. Grave de Peralta Menendez, M.M. Murray, C.M. Michel, R. Martuzzi, and S.L. Gonzalez Andino. Electrical neuroimaging based on biophysical constraints. *NeuroImage*, 21:527–539, 2004.

[71] A. Delorme and S. Makeig. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods*, 134:9–21, 2004.

[72] A. Delorme, J. Palmer, J. Onton, R. Oostenveld, and S. Makeig. Independent EEG sources are dipolar. *PLoS One*, 7:e30135, 2012.

[73] A. Delorme, M. Westerfield, and S. Makeig. Medial prefrontal theta bursts precede rapid motor responses during visual selective attention. *The Journal of Neuroscience*, 27:11949–11959, 2007.

[74] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society*, 39:1–38, 1977.

[75] T. Dikanev, D. Smirnov, R. Wennberg, J.L. Perez Velazquex, and B. Bezruchko. EEG nonstationarity during intracranially recorded seizures: statistical and dynamical analysis. *Clinical Neurophysiology*, 116:1796–1807, 2005.

[76] M. Ding, S.L. Bressler, W. Yang, and H. Liang. Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: data preprocessing, model validation, and variability assessment. *Biol. Cybern.*, 83:35–45, 2000.

[77] M. Dyrholm, S. Makeig, and L. K. Hansen. Model selection for convolutive ICA with an application to spatiotemporal analysis of EEG. *Neural Computation*, 19:934–955, 2007.

[78] B. Efron. Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, 7:1–26, 1979.

[79] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani. Least angle regression. *The Annals of Statistics*, 32:407–499, 2004.

[80] W. Einthoven. The string galvanometer and the human electrocardiogram. In *KNAW Proceedings*, volume 6, pages 107–115, 1903.

[81] R. Elul. Randomness and synchrony in the generation of the electroencephalogram. In H. Petsche and M.A.B. Brazier, editors, *Synchronization of EEG Activity in Epilepsies*. Springer-Verlag, 1972.

[82] A. K. Engel and P. Fries. Beta-band oscillations – signalling the status quo? *Current Opinion in Neurobiology*, 20:156–165, 2010.

[83] B.A. Eriksen and C.W. Eriksen. Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept Psychophys*, 16:143–149, 1974.

[84] J.J. Ermer, J.C. Mosher, S. Baillet, and R.M. Leahy. Rapidly recomputable EEG forward models for realistic head shapes. *Phys Med Biol*, 46:1265–1281, 2001.

[85] M. C. Espana-Boquera and A. Puerta-Notario. Noise effects in injection locked laser simulation: Phase jumps and associated spectral components. *Electronics Letters*, 32:818–819, 1996.

[86] D.J. Felleman and D.C. Van Essen. Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex*, 1:1–47, 1991.

[87] R.P. Feynman. *Statistical Mechanics: A Set of Lectures*. W.A. Benjaminm Inc, 1972.

[88] L.H. Finkel and G.M. Edelman. Integration of distributed cortical systems by reentry: A computer simulation of interactive functionally segregated visual areas. *J Neurosci*, 9:3188–3208, 1989.

[89] N. I. Fisher. *Statistical Analysis of Circular Data*. Cambridge University Press, 1995.

[90] A. Forbes and C. Thacher. Amplification of action currents with the electron tube in recording with the string galvanometer. *Amer J Physiol*, 52:409–471, 1920.

[91] A. Fourment, L. Jami, J. Calvet, and J. Scherrer. Comparison of the EEG recorded from the scalp with the elementary activity of radial cortical dipoles. *Electroencephalogr Clin Neurophysiol*, 19:217–229, 1965.

[92] A.M. Fraser and H.L. Swinney. Independent coordinates for strange attractors from mutual information. *Phys Rev A*, 33:1134–1140, 1986.

[93] W.J. Freeman, B.C. Burke, and M.D. Holmes. Aperiodic phase re-setting in scalp EEG of beta-gamma oscillations by state transitions at alpha-theta rates. *Hum Brain Mapp*, 19:248–272, 2003.

[94] W.J. Freeman and L.J. Rogers. Fine temporal resolution of analytic phase reveals episodic synchronization by state transitions in gamma EEGs. *Journal of Neurophysiology*, 87:937–945, 2001.

[95] J. H. Friedman. Recent advances in predictive (machine) learning. *Journal of Classification*, 23:175–197, 2006.

[96] K. Friston. The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11.2:127–138, 2010.

[97] K. Friston, L. Harrison, J. Daunizeau, S. Kiebel, C. Phillips, N. Trujillo-Barreto, R. Henson, G. Fladin, and J. Mattout. Multiple sparse priors for the M/EEG inverse problem. *Neuroimage*, 39:1104–1120, 2008.

[98] K. Friston, L. Harrison, and W. Penny. Dynamic causal modelling. *NeuroImage*, 19:1273–1302, 2003.

[99] K.J. Friston. Functional and effective connectivity in neuroimaging: A synthesis. *Hum Brain Mapp*, 2:56–78, 1994.

[100] K.J. Friston, W. Penny, C. Phillips, S. Kiebel, G. Hinton, and J. Ashburner. Classical and bayesian inference in neuroimaging: Theory. *Neuroimage*, 16:465–483, 2002.

[101] D. Gabor. Theory of communication. *J Inst Electr Eng*, 93:429–457, 1946.

[102] J. Galambos. On the distribution of the maximum of random variables. *The Annals of Mathematical Statistics*, 43:516–521, 1972.

[103] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.

[104] D.B. Geselowitz. On bioelectric potentials in an inhomogeneous volume conductor. *Biophysical Journal*, 7, 1967.

[105] K. Ghosh, S. R. Jammalamadaka, and M. Vasudaven. Change-point problems for the von Mises distribution. *Journal of Applied Statistics*, 26:423–434, 1999.

[106] D. Giannitrapani, V.T. Rast, and B.J. Shulhafer. Multiple channel direct digital recording of EEG data. *Behavioural Science*, 16:239–243, 1971.

[107] F.A. Gibbs, H. Davis, and W.G. Lennox. The electro-encephalogram in epilepsy and conditions of impaired conciousness. *Archives of Neurology and Phychiatry*, 36:1133–1148, 1935.

[108] J.W. Gibbs. On the equilibrium of heterogeneous substances. *Transactions of the Connecticut Academy*, 3:108–248, 343–524, 1877.

[109] D. Gibert and J-L. Le Mou el. Inversion of polar motion data: Chandler wobble, phase jumps, and geomagnetic jerks. *Journal of Geophysical Research*, 113:B10405, 2008.

[110] A. Glass and S.F. Gurney. A percent time alpha computer for the electroencephalograph. *Wld Med Electron Instrum*, 2:206–210, 1964.

[111] G. H. Golub, M. Heath, and G. Wahba. Genearlized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21:215–223, 1979.

[112] G. Gomez-Herrero, M. Atienza, K. Egiazarian, and J. L. Cantero. Measuring directed coupling between EEG sources. *NeuroImage*, 43:497–508, 2008.

[113] N.R. Goodman. Measuring amplitude and phase. *J Franklin Inst*, 270:437–450, 1960.

[114] I.F. Gorodnitsky, J.S. George, and B.D. Rao. Neuromagnetic source imaging with FOCUSS: a recursive weighted minimum norm algorithm. *Electroencephalogr Clin Neurophysiol*, 95:231–251, 1995.

[115] I. Grabovsky and L. Horváth. Change-point detection in angular data. *Ann. Inst. Statist. Math.*, 53:552–566, 2001.

[116] C.W.J. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37:424–438, 1969.

[117] A.M. Grass. The electroencephalography heritage until 1960. *American Journal of EEG Technology*, 24, 1984.

[118] G. Gratton, M.G.H. Coles, and E. Donchin. A new method for off-line removal of ocular artifact. *Electroencephalogr Clin Neurophysiol*, 55:468–484, 1983.

[119] C.M. Gray. The temporal correlation hypothesis of visual feature integration: Still alive and well. *Neuron*, 24:31–47, 1999.

[120] C.M. Gray, P. Konig, A.K. Engel, and W. Singer. Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, 338:334–337, 1989.

[121] R. Grech, T. Cassar, J. Muscat, K.P. Camilleri, S.G. Fabri, M. Zervakis, P. Xanthopoulos, V. Sakkalis, and B. Vanrumste. Review on solving the inverse problem in EEG source analysis. *J Neuroeng Rehabil*, 5, 2008.

[122] S. W. Greenhouse and S. Geisser. On methods in the analysis of profile data. *Psychometrika*, 24:95–112, 1959.

[123] A. N. Grigorenko, P.I. Nikitin, and A. V. Kabashin. Phase jumps and interferometric surface plasmon resonance imaging. *Applied Physics Letters*, 75:3917–3919, 1999.

[124] R.C. Gur, B.I. Turetsky, M. Matsui, M. Yan, W. Bilker, P. Hughett, and R.E. Gur. Sex differences in brain grey and white matter in healthy young adults: Correlations with cognitive performance. *J Neurosci*, 19:4065–4072, 1999.

[125] D. Hagemann, E. Naumann, and J.F. Thayer. The quest for the EEG reference revisited: A glance from brain asymmetry research. *Psychophysiology*, 38:847–857, 2001.

[126] P. Hall, J. L. Horowitz, and B. Jing. On blocking rules for the bootstrap with dependent data. *Biometrika*, 82:561–574, 1995.

[127] A.M. Halliday. Cortical EPs in man: Clinical observations. In E. Callaway, P. Tueting, and S.H. Koslow, editors, *Spatial Contrast: Report of a Workship*. London; Academic Press, 1977.

[128] A.M. Halliday, E. Halliday, A. Kriss, W.I. McDonald, and J. Mushin. The patter-evoked potential in compression of the anterior visual pathways. *Brain*, 99:357–374, 1976.

[129] M. Hamalainen, R. Hari, R.J. Ilmoniemi, J. Knuutila, and O.V. Lounasmaa. Magnetoencephalography - theory, instrumentation, and applications to noninvasive studies of the working brain. *Review of Modern Physics*, 65:413–504, 1993.

[130] M.S. Hämäläinen and R.J. Ilmoniemi. Interpreting measured magnetic fields of the brain: estimates of current distributions. Technical Report TKK-F-A559, Helsinki University of Technology, 1984.

[131] M.S. Hämäläinen and R.J. Ilmoniemi. Interpreting magnetic fields of the brain: minimum norm estimates. *Med Biol Eng Comput*, 32:35–42, 1994.

[132] M. Härle, B. S. Rochstroh, A. Keil, C. Wienbruch, and T. R. Elbert. Mapping the brain's orchestration during speech comprehension: Task-specific facilitation of regional synchrony in neural networks. *Biomed Central Neuroscience*, 5:40, 2004.

[133] S. Haufe, V. Nikulin, K-R. Müller, A. Ziehe, and G. Nolte. Estimating vector fields using sparse basis field expansions. In D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, editors, *Advances in Neural Information Processing Systems 21*, pages 617–624. MIT Press, Cambridge, 2009.

[134] S. Haufe, V. V. Nikulin, K. R. Müller, and G. Nolte. Combining sparsity and rotational invariance in EEG/MEG source reconstruction. *Neuroimage*, 42:726–738, 2008.

[135] S. Haufe, V.V. Nikulin, K-R. Müller, and G. Nolte. A critical assessment of connectivity measures for EEG data: A simulation study. *NeuroImage*, 64:120–133, 2013.

[136] S. Haufe, R. Tomioka, T. Dickhaus, C. Sannelli, B. Blankertz, G. Nolte, and K-R. Müller. Large-scale EEG/MEG source localization with spatial flexibility. *Neuroimage*, 54:851–859, 2011.

[137] S. Haufe, R. Tomioka, G. Nolte, K-R. Müller, and M. Kawanabe. Modeling sparse connectivity between underlying brain sources for EEG/MEG. *IEEE Transactions on Biomedical Engineering*, 57:1954–1963, 2010.

[138] D.O. Hebb. *The organization of behavior: A neurophysiological theory.* New York:Wiley, 1949.

[139] C.J. Henderson, S.R. Butler, and A. Glass. The localization of equivalent dipoles of EEG sources by the application of electrical field theory. *Electroencephalography and Clincal Neurophysiology*, 39:117–130, 1975.

[140] W. Hesse, E. Moller, M. Arnold, and B. Schack. The use of time-variant EEG granger causality for inspecting directed interdependencies of neural assemblies. *J Neurosci Methods*, 124:27–44, 2003.

[141] S. Holm. A simple sequentially rejected multiple test proceedure. *Scand J Statist*, 6:65–70, 1979.

[142] M. Huskova and C. Kirch. Bootstrapping confidence intervals for the change-point of time series. *Journal of Time Series Analysis*, 29:947–972, 2008.

[143] A. Hyvarinen and E. Oja. Independent component analysis: Algorithms and applications. *Neural Networks*, 13:411–430, 2000.

[144] A.A. Ioannides, J.P.R. Bolton, and C.J.S. Clark. Continuous probabilistic solutions to the biomagnetic inverse problem. *Inverse Problems*, 6:523–542, 1990.

[145] J. Ito, A. R. Nikolaev, and C. V. Leeuwen. Dynamics of spontaneous transitions between global brain states. *Human Brain Mapping*, 28:904–913, 2007.

[146] E. M. Izhikevich. Synchronization. In *Dynamical Systems in Neuroscience: The geometry of excitability and bursting*, chapter 10, pages 443–505. MIT Press, 2006.

[147] H.H. Jasper. Report on the committee on methods of clinical examination in electroencephalography. *Electroencephalogr Clin Neurophysiol*, 10:370–375, 1958.

[148] J. Jeong, J.C. Gore, and B.S. Peterson. Mutual information analysis of the EEG in patients with alzheimer's disease. *Clincal Neurophysiology*, 112:827–835, 2001.

[149] K. Jerbi, J-P. Lachaux, K. N'Diaye, D. Pantazis, R. M. Leahy, L. Garnero, and S. Baillet. Coherent neural representation of hand speed in humans revealed by MEG imaging. *PNAS*, 104:7676–7681, 2007.

[150] D.L. Jewett and J.S. Williston. Auditory-evoked far fields averaged from the scalp of humans. *Brain*, 94:681–696, 1971.

[151] E.R. John. From synchronous neural discharges to subjective awareness? In S. Laureys, editor, *Progress in Brain Research*, volume 150, chapter 11. Elsevier B.V., 2005.

[152] S.C. Jun, J.S. George, S.M. Plis, D.M. Ranken, D.M. Schmidt, and C.C. Wood. Improving source detection and seperation in a spatiotemporal bayesian inference dipoles analysis. *Physics in Medicine and Biology*, 51:2395–2414, 2006.

[153] C. Jutten and J. Herault. Blind seperation of sources, part I: An adaptive algorithm based on neuromimetic architecture. *Signal Processing*, 24:1–10, 1991.

[154] M. Kaminski, M. Ding, W.A. Truccolo, and S.L. Bressler. Evaluating causal relationships in neural systems: Granger causality, directed transfer function and statistical assessment of significance. *Biol Cybern*, 85:145–147, 2001.

[155] A.Y. Kaplan, A.A. Fingelkurts, A.A. Fingelkurts, S.V. Borisov, and B.S. Darkhovsky. Nonstationary nature of the brain activity as revealed by EEG/MEG: Methodological, practical and conceptual challenges. *Signal Processing*, 85:2190–2212, 2005.

[156] R.N. Kavanagh, T.M. Darcey, D. Lehmann, and D.H. Fender. Evaluation of methods for three-dimensional localization of electrical sources in the human brain. *IEEE Trans. Biomed. Engineering*, 25:421–429, 1978.

[157] S. J. Kiebel, M. I. Garrido, R. J. Moran, and K. J. Friston. Dynamic causal modelling for EEG and MEG. *Cognitive Neurodynamics*, 2:121–136, 2008.

[158] S.J. Kiebel and K.J. Friston. Statistical parametric mapping for event-related potentials (ii): a hierarchical temporal model. *NeuroImage*, 22:503–520, 2004.

[159] T. Kiemel, K.M. Gormley, L. Guan, T.L. Williams, and A.H. Cohen. Estimating strength and direction of functional coupling in the lamprey spinal cord. *J Comput Neurosci*, 15:233–245, 2003.

[160] S. Kim, D. Putrino, S. Ghosh, and E.N. Brown. A granger causality measure for point process models of ensemble neural spiking activity. *PLoS Comput Biol*, 7:1–13, 2011.

[161] C. Kirch. Block permutation principles for the change analysis of dependent data. *Journal of Statistical Planning and Inference*, 137:2453–2474, 2007.

[162] C. Kirch. Bootstrapping sequential change-point tests. *Sequential Analysis: Design Methods and Applications*, 27:330–349, 2008.

[163] G. Kobel and C. Hummel. Cerebral chemosensory evoked potentials elicited by chemical stimulation of the human olfactory and respiratory nasal mucosa. *Electroencephalogr Clin Neurophysiol*, 71:241–250, 1988.

[164] T. Koeda, M. Knyazeva, C. Njiokiktjien, E.J. Jonkman, L. De Sonneville, and V. Vildavsky. The EEG in acallosal children. coherence values in the resting state: left hemisphere compensatory mechanism. *Electroencephalogr Clin Neurophysiol*, 95:397–407, 1995.

[165] H. R. Künsch. The jackknife and the bootstrap for general stationary observations. *Ann. Statist.*, 17:1217–1241, 1989.

[166] J-P. Lachaux, E. Rodriguez, J. Martinerie, and F.J. Varela. Measuring phase synchrony in brain signals. *Hum Brain Mapp*, 8:194–208, 1999.

[167] C.L. Lackner, W.J. Marshall, D.L. Santesso, J. Dywan, T. Wade, and S.J. Segalowitz. Adolescent anxiety and aggression can be differentially predicted by electrocortical phase reset variables. *Brain and Cognition*, (In Press), 2013.

[168] S. N. Lahiri. Theoretical comparisons of block bootstrap methods. *The Annals of Statistics*, 27:386–404, 1999.

[169] K.S. Lashley and G. Clark. The cytoarchitecture of the cerebral cortex of ateles: A critical examination of architectonic studies. *J Comparative Neurology*, 85:223–305, 1946.

[170] P.W. Lee, Z.J. Wang, and M.J. McKeown. Mutual information based relevance network analysis: A parkinson's disease study. *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 497–500, 2008.

[171] T. W. Lee, M. Girolami, and T. J. Sejnowski. Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Computation*, 11:417–441, 1999.

[172] S. Lemm, B. Blankertz, T. Dickhaus, and K. R. Müller. Introduction to machine learning for brain imaging. *Neuroimage*, 56:387–399, 2011.

[173] C. Li and H. Jasper. Microelectrode studies of the electrical activity of the cerebral cortex in the cat. *J Physiol*, 121:117–140, 1953.

[174] Y. Li, Z. Ma, W. Lu, and Y. Li. Automatic removal of the eye blink artifact from EEG using an ICA-based template matching approach. *Physiology Measures*, 27:425–436, 2006.

[175] W. Liao, Z. Zhang, Z. Pan, D. Mantini, J. Ding, X. Duan, C. Luo, G. Lu, and H. Chen. Altered functional connectivity and small-world in mesial temporal lobe epilepsy. *PLoS One*, 5:e8525, 2010.

[176] Y. Liu, M. Liang, Y. Zhou, Y. He, Y. Hao, M. Song, C. Yu, H. Liu, Z. Liu, and T. Jiang. Disrupted small-world networks in schizophrenia. *Brain*, 131:945–961, 2008.

[177] M. Livingston and D. Hubel. Segregation of form, color, movement, and depth: Anatomy,physiology, and perception. *Science*, 240:740–749, 1988.

[178] R.R. Llinás and U. Ribary. Temporal conjunction in thalamocortical transactions. In H.H. Jasper, L. Descarries, V.F. Castellucci, and S. Rossignol, editors, *Consciousness: At the Frontiers of Neuroscience, Advances in Neurology*, volume 77. Lippincotti-Raven Publishers, Philadelphia, 1998.

[179] F. Lombard. The change-point problem for angular data: A nonparametric approach. *Technometrics*, 28:391–397, 1986.

[180] S. Makeig, T-P. Jung, A. J. Bells, D. Ghahremani, and T. J. Sejnowski. Blind separation of auditory event-related brain responses into independent components. *Proc. Natl. Acad. Sci. USA*, 94:10979–10984, 1997.

[181] S. Makeig, M. Westerfield, T.-P. Jung, S. Enghoff, J. Townsend, E. Courchesne, and T.J. Sejnowski. Dynamic brain sources of visual evoked responses. *Science*, 295:690–694, 2002.

[182] N. J. I. Mars, F. H. Lopes Da Silva, K. van Hulten, and J. G. Lommen. Computer assisted analysis of EEGs during seizures; location of an epileptogenic area. *Electroencephalogr Clin Neurophysiol*, 43:575, 1977.

[183] N.J.I. Mars, P.M. Thompson, and R.J. Wilkus. Spread of epileptic seizure activity in humans. *Epilepsia*, 26:85–94, 1985.

[184] N. F. Martínez and R. García. Measuring phase shifts and energy dissipation with amplitude modulation atomic force microscopy. *Nanotechnology*, 17:S167–S172, 2006.

[185] J. Mattout, C. Phillips, W.D. Penny, M.D. Rugg, and K.J. Friston. Meg source localization under multiple constraints: An extended bayesian framework. *Neuroimage*, 30:753–767, 2006.

[186] C. A. McGilchrist and K. D. Woodyer. Note on a distribution-free cusum technique. *Technometrics*, 17:321–325, 1975.

[187] S. Mead, Francis J.P. Ebling, E. S. Maywood, T. Humbly, J. Herbert, and M. H. Hastings. A nonphotic stimulus causes instantaneous phase advances of the light-entrainable circadian oscillator of the syrian hamster but does not induce the expression of c-fos in the suprachiasmatic nuclei. *J. Neurosci.*, 12:2516–2522, 1992.

[188] J.d.R. Millán, R. Rupp, G.R. Müller-Putz, R. Murray-Smith, C. Giugliemma, M. Tangermann, C. Vidaurre, F. Cincotti, A. Kübler, R. Leeb, C. Neuper, K-R. Müller, and D. Mattia. Combining brain-computer interfaces and assistive technologies: State-of-the-art and challenges. *Frontiers in Neuroscience*, 4:161, 2010.

[189] E. Moller, B. Schack, M. Arnold, and H. Witte. Instantaneous multivariate EEG coherence analysis by means of adaptive high-dimensional auto-regressive models. *J Neurosci Methods*, 105:143–158, 2001.

[190] F. Mormann, K. Lehnertz, P. David, and C.E. Elger. Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients. *Physica D*, 144:358–369, 2000.

[191] J.C. Mosher and R.M. Leahy. Source localization using recursively applied and projected (RAP) MUSIC. *IEEE Transactions on Signal Processing*, 47:332–340, 1999.

[192] J.C. Mosher, R.M. Leahy, and P.S. Lewis. EEG and MEG: Forward solutions for inverse problems. *IEEE Trans on Biomdeical Engineering*, 46:245–259, 1999.

[193] J.C. Mosher, P.S. Lewis, and R.M. Leahy. Multiple dipole modeling and localization from spatio-temporal MEG data. *IEEE Transactions on Biomedical Engineering*, 39:541–557, 1992.

[194] M. Murias, S. J. Webb, J. Greenson, and G. Dawson. Resting state cortical connectivity reflected in EEG coherence in individuals with autism. *Biol. Psychiatry*, 62:270–273, 2007.

[195] S.H. Na, S. Jin, and S.Y. Kim. The effects of total sleep deprivation on brain functional organization: Mutual information analysis of waking human EEG. *Int. Journal of Psychophysiology*, 62:238–242, 2006.

[196] S.H. Na, S. Jin, S.Y. Kim, and B. Ham. EEG in schizophrenic patients: mutual information analysis. *Clinical Neurophysiology*, 113:1954–1960, 2002.

[197] R. Näätänen and T. Picton. The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiol*, 24:375–425, 1987.

[198] R.M. Neal. *Bayesian Learning for Neural Networks*. Springer-Verlag, 1996.

[199] P.W. Nicholson. Specific impedance of cerebral white matter. *Experimental Neurology*, 13:386–401, 1965.

[200] E. Niedermeyer and F. Lopes da Silva. *Electroencephalography : Basic Principles, Clinical Applications and Related Fields*. Urban and Schwarzenberg, second edition, 1987.

[201] G. Nolte, O. Bai, L. Wheaton, Z. Mari, S. Vorbach, and M. Hallett. Identifying true brain interaction from EEG data using the imaginary part of coherency. *Clin Neurophysiol*, 115:2292–2307, 2004.

[202] G. Nolte, A. Ziehe, V. Nikulin, A. Schlögl, N. Krämer, T. Brismar, and K-R. Müller. Robustly estimating the flow direction of information in complex physical systems. *Phys Rev Lett*, 100:234101, 2008.

[203] A. Nummenmaa, T. Auranen, M.S. Hamalainen, I.P. Jaaskelainen, J. Lampinen, M. Sams, and A. Vehtari. Hierarchical bayesian estimates of distributed meg sources: theoretical aspects and comparison of variational and mcmc methods. *Neuroimage*, 35:669–685, 2007.

[204] P.L. Nunez, R.B. Silberstein, Z. Shi, M.R. Carpenter, R. Srinivasan, D.M. Tucker, S.M. Doran, P.J. Cadusch, and R.S. Wijesinghe. EEG coherency II: experimental comparisons of multiple measures. *Clinical neurophysiology*, 110:469–486, 1999.

[205] P.L. Nunez, R. Srinivasan, A.F. Westdorp, R.S. Wijesinghe, D.M. Tucker, R.B. Silberstein, and P.J. Cadusch. EEG coherency I: statistics, reference electrode, volume conduction, laplacians, cortical imaging, and interpretation at multiple scales. *Electroencephalogr Clin Neurophysiol*, 103:499–515, 1997.

[206] H. Nyquist. Certain factors affecting telegraph speed. *Journal of the AIEE*, 43:124–146, 1924.

[207] J. Onton, A. Delorme, and S. Makeig. Frontal midline EEG dynamics during working memory. *NeuroImage*, 27:341–356, 2005.

[208] R. Oostenveld, P. Fries, E. Maris, and J. Schoffelen. Fieldtrip: Open source software for advanced analysis of MEG, EEG, and inverse electrophysiological data. *Computation Intelligence and Neuroscience*, 2011, 2011.

[209] G. V. Osipov, A. S. Pikovsky, M. G. Rosenblum, and J. Kurths. Phase synchronization effects in a lattice of nonidentical rössler oscillators. *Physical Review E*, 55:2353–2361, 1997.

[210] E. S. Page. Continuous inspection schemes. *Biometrika*, 41:100–115, 1945.

[211] E. S. Page. A test for a change in a parameter occuring at an unknown point. *Biometrika*, 42:523–527, 1955.

[212] S. Palit. *Prinicples of Electricity and Magnetism*. Alpha Science International Ltd, 2005.

[213] M. Palus, V. Komarek, Z. Hrncir, and K. Sterbova. Synchronization as adjustment of information rates: Detection from bivariate time series. *Phys Rev E*, 63:1–6, 2001.

[214] R.D. Pascual-Marqui. Review of methods for solving the EEG inverse problem. *International Journal of Bioelectromagnetism*, 1:75–86, 1999.

[215] R.D. Pascual-Marqui. Standardized low resolution brain electromagnetic tomography (sLORETA): technical details. *Methods & Findings in Experimental & Clincal Pharmacology*, 24D:5–12, 2002.

[216] R.D. Pascual-Marqui, C.M. Michel, and D. Lehmann. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *Int. Journal of Psychophysiology*, 18:49–65, 1994.

[217] A. N. Pettitt. A non-parametric approach to the change-point problem. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28:126–135, 1979.

[218] G. Pfurtscheller. Central beta rhythm during sensorimotor activities in man. *Electroencephalography and Clinical Neurophysiology*, 51:253–264, 1981.

[219] C. Phillips, J. Mattout, M.D. Rugg, P. Maquet, and K.J. Friston. An empirical bayesian solution to the source reconstruction problem in EEG. *NeuroImage*, 24:997–1011, 2005.

[220] J.W. Phillips, R.M. Leahy, and J.C. Mosher. Meg-based imaging of focal neuronal current source. *IEEE Transactions on Medical Imaging*, 16:338–348, 1997.

[221] A. Pikovsky, M. Rosenblum, and J. Knuths. *Synchronization: A universal concept in nonlinear sciences*. Cambridge Univ Press, New York, 2003.

[222] A.S. Pikovsky. Phase synchronization of chaotic oscillators by a periodic external field. *Sov. J. Commun. Technol. Electron.*, 30, 1985.

[223] A.S. Pikovsky, M.G. Rosenblum, G.V. Osipov, and J. Kurths. Phase synchronization of chaotic oscillators by external driving. *Physica D*, 104:219–238, 1997.

[224] D. Politis and J. P. Romano. A circular block resampling procedure for stationary data. In R. Lepage and L. Billard, editors, *Exploring the Limits of Bootstrap*, pages 263–270. Wiley, New York, 1992.

[225] H. M. Praetorius, G. Bodenstein, and O.D. Cruetzfeldt. Adaptive segmentation of EEG records: A new approach to automatic EEG analysis. *Electroencephalogr Clin Neurophysiol.*, 42:84–94, 1977.

[226] R.Q. Quiroga, A. Kraskov, T. Kreuz, and P. Grassberger. Performance of different synchronization measures in real data: A case study on electroencephalographic signals. *Phys Rev E*, 65:041903, 2002.

[227] M. Le Van Quyen, J. Foucher, J-P. Lachaux, E. Rodriguez, A. Lutz, J. Martinerie, and F.J. Varela. Comparison of hilbert transform and wavelet methods for the analysis of neural synchrony. *J Neurosci Methods*, 111:83–98, 2001.

[228] P. Ramanand, M.C. Bruce, and E.N. Bruce. Mutual information analysis of EEG signals indicates age-related changes in cortical interdependence during sleep in middle aged versus elderly women. *Journal of Clinical Neurophysiology*, 27:274–284, 2010.

[229] P. Rappelsberger. The reference problem and mapping of coherene: A simulation study. *Brain Topography*, 2:63–72, 1989.

[230] A. Rémond. The importance of topographic data in EEG phenomena and an electric model to reproduce them. *Electroencephlogr Clin Neurophysiol Suppl*, 27, 1968.

[231] S.J. Roberts and W. Penny. Variational bayes for generalized autoregressive models. *IEEE Trans Biomed Eng*, 50:2245–2257, 2002.

[232] E. Rodriguez, N. George, J. Lachaux, J. Martinerie, B. Renault, and F.J. Varela. Perception's shadow: long-distance synchronization of human brain activity. *Nature*, 397:430–433, 1999.

[233] M.G. Rosenblum, L. Cimponeriu, A. Bezerianos, A. Patzak, and R. Mrowka. Identification of coupling direction: Application to cardiorespiratory interation. *Phys Rev E*, 65:041909, 2002.

[234] M.G. Rosenblum and A.S. Pikovsky. Detecting direction of coupling in interacting oscillators. *Phys Rev E*, 64:045202, 2001.

[235] M.G. Rosenblum, A.S. Pikovsky, and J. Kurths. Phase synchronization of chaotic oscillators. *Phys Rev Letters*, 76:1804–1807, 1996.

[236] O. E. Rössler. An equation for continuous chaos. *Physics Letters*, 57A:397–398, 1976.

[237] M. Roth, J. Shaw, and J. Green. The form, voltage distribution and physiological significance of the k-complex. *Electroencephalogr Clin Neurophysiol*, 8:385–402, 1956.

[238] M.R. Rueda, M.I. Posner, and M.K. Rothbart. The development of executive attention: Contributions to the emergence of self-regulation. *Dev Neuropsychol*, 28:573–594, 2005.

[239] S. Rush and D.A. Driscoll. EEG electrode sensitivity - an application of reciprocity. *IEEE Trans Biomed Engineering*, BME-16:15–22, 1969.

[240] G.S. Russell, R. Srinivasan, and D.M. Tucker. Bayesian estimates of error bounds for EEG source imaging. *IEEE Transactions on Biomedical Imaging*, 17:1084–1089, 1998.

[241] Y. Saito and H. Harashima. Tracking information within multichannel EEG record: Causal analysis in EEG. In N. Yamaguchi and K. Fujisawa, editors, *Recent Advances in EEG and MEG data processing*, pages 133–146. Elsevier, 1981.

[242] R. Sanchez, D.E. Newman, and B.A. Carreras. Waiting-time statistics of self-organized-criticality systems. *Phys Rev Lett*, 88:1–4, 2002.

[243] M-S. Sato, T. Yoshioka, S. Kajihara, K. Toyama, N. Goda, K. Doya, and M. Kawato. Hierarchical bayesian estimation for meg inverse problem. *Neuroimage*, 23:806–826, 2004.

[244] L.K. Saul, T. Jaakkola, and M.I. Jordan. Mean field theory for sigmoid belief networks. *Journal of Artificial Intelligence Research*, 4:61–76, 1996.

[245] B. Schelter, M. Winterhalder, M. Eichler, M. Peifer, B. Hellwig, B. Guschlbauer, C.H. Lucking, R. Dahlhaus, and J. Timmer. Testing for directed influences amoung neural signals using partial directed coherence. *J Neurosci Methods*, 2005.

[246] M. Scherg and D. Von Cramon. Evoked dipole source potentials of the human auditory cortex. *Electroencephalography and Clincal Neurophysiology*, 65:344–360, 1986.

[247] D.M. Schmidt, J.S. George, and C.C. Wood. Bayesian inference applied to the electromagnetic inverse problem. *Hum Brain Mapp*, 7:195–212, 1999.

[248] M.R. Schneider. A multistage process for computing virtual dipolar sources of EEG discharges from surface information. *IEEE Trans Biomedical Engineering*, BME-19:1–12, 1972.

[249] J. Schoffelen and J. Gross. Source connectivity analysis with MEG and EEG. *Hum Brain Mapp*, 30:1857–1865, 2009.

[250] T. Schreiber. Measuring information transfer. *Phys Rev Letters*, 85:461–464, 2000.

[251] H.P. Schwan and C.F. Kay. Specific resistance of body tissues. *Circulation Research*, 4:664–670, 1956.

[252] A. Sen and M. S. Srivastava. On tests for detecting change in mean. *The Annals of Statistics*, 3:98–108, 1975.

[253] C. Shannon. Communication in the presence of noise. *Proceedings of the IRE*, 37:10–21, 1949.

[254] C.E. Shannon. A mathematical theory of communication. *The Bell Systems Technical Journal*, 27:379–423, 623–656, 1948.

[255] D.W. Shattuck, S.R. Sandor-Leahy, K.A. Schaper, D.A. Rottenberg, and R.M. Leahy. Magnetic resonance image tissue classification using a partial volume model. *NeuroImage*, 13:856–876, 2001.

[256] R.L. Silton, W. Heller, D.N. Towers, A.S. Engels, J.M. Spielberg, J.C. Edgar, S.M. Sass, J.L. Stewart, B.P. Sutton, M.T. Banich, and G.A. Miller. The time course of activity in dorsalateral prefrontal cortex and anterior cingulate cortex during top-down attention control. *NeuroImage*, 50:1292–1302, 2010.

[257] M. Sinn, A. Ghodsi, and K. Keller. Detecting change-points in time series by maximum mean discrepancy of ordinal pattern distributions. In *Uncertainty in Artificial Intelligence: Proceedings of the Twenty-Eighth Conference*, 2012.

[258] B. Sklar, J. Hanley, and W. W. Simmons. An eeg experiment aimed toward identifying dyslexic children. *Nature*, 240:414–416, 1972.

[259] D.A. Smirnov and B.P. Bezruchko. Estimation of interaction strength and direction from short and noisy time series. *Phys Rev E*, 68:046209, 2003.

[260] O. Sporns, D.R. Chialvo, M. Kaiser, and C.C. Hilgetag. Organization, development and function of complex brain networks. *TRENDS in Cognitive Science*, 8:418–425, 2004.

[261] O. Sporns, G. Tononi, and G.M. Edelman. Modeling perceptual grouping and figure-ground segregation by means of active reentry connections. *Proc Natl Acad Sci U.S.A*, 88:129–133, 1991.

[262] O. Sporns, G. Tononi, and G.M. Edelman. Theoretical neuroanatomy: Relating anatomical and functional connectivity in graphs and cortical connection matrices. *Cereb Cortex*, 10:127–141, 2000.

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

[264] C. J. Stam, G. Nolte, and A. Daffertshofer. Phase lag index: Assessment of functional connectivity from multi-channel EEG and MEG with diminished bias from common sources. *Human Brain Mapping*, 28:1178–1193, 2007.

[265] M. Steriade. Cellular substrates of brain rhythms. In E. Niedermeyer and F. Lopes da Silva, editors, *Electroencephalography*. Williams and Wilkins, Baltimore, 1995.

[266] G. Sugihara, R. May, H. Ye, C. Hsieh, E. Deyle, M. Fogarty, and S. Munch. Detecting causality in complex ecosystems. *Science*, 338:496–500, 2012.

[267] S. Sutton, M. Braren, J. Zubin, and E.R. John. Evoked-potential correlates of stimulus uncertainty. *Science*, 150:1187–1188, 1965.

[268] D.Y. Takahashi, L.A. Baccala, and K. Sameshime. Connectivity inference between neural structures via partial directed coherence. *J Appl Stat*, 34:1255–1269, 2007.

[269] M. T. Taner, F. Koehler, and R. E. Sheriff. Complex seismic trace analysis. *Geophysics*, 44:1041–1063, 1979.

[270] P. Tass, M.G. Rosenblum, J. Weule, J. Kurths, A. Pikovsky, J. Volkmann, A. Schnitzler, and H.J. Freund. Detection of n:m phase locking from noisy data: Application to magnetoencephalography. *Phys Rev Lett*, 81:3291–3294, 1998.

[271] M. ten Caat, M. M. Lorist, E. Bezdan, J. B.T.M. Roerdink, and N. M. Maurits. High-density EEG coherence analysis using functional units applied to mental fatigue. *J. Neuroscience Methods*, 171:271–278, 2008.

[272] R. Thatcher, D. North, and C. Biver. EEG and intelligence: Univariate and multivariate comparisons between EEG coherence, EEG phase delay and power. *Clin Neurophysiol*, 116, 2005.

[273] R. W. Thatcher, D. M. North, J. Neubrander, and C. J. Biver. Autism and EEG phase reset: Deficient GABA mediated inhibition in thalamo-cortical circuits. *Developmental Neuropsychology*, 34:780–800, 2009.

[274] R.W. Thatcher. Cyclic cortical reorganization during early childhood. *Brain Cogn*, 20:24–50, 1992.

[275] R.W. Thatcher. Coherence, phase differences, phase shift and phase lock in EEG/ERP analysis (in progress), 2010.

[276] R.W. Thatcher, P.J. Krause, and M. Hrybyk. Cortico-cortical associations and EEG coherence: a two-compartmental model. *Electroencephalogr Clin Neurophysiol*, 64:123–143, 1986.

[277] R.W. Thatcher, D.M. North, and C.J. Biver. Self-organized criticality and the development of EEG phase reset. *Hum Brain Mapp*, 30:553–574, 2009.

[278] R. Tibshirani. Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society. Series B*, 58:267–288, 1996.

[279] R. Tibshirani, M. Saunders, S. Rosset, J. Zhu, and K. Knight. Sparsity and smoothness via the fused lasso. *Journal of the Royal Statistical Society. Series B*, 67:91–108, 2005.

[280] G. Tononi, O. Sporns, and G.M. Edelman. Reentry and the problem of integrating multiple cortical areas: Simulation of dynamic integration in the visual system. *Cereb Cortex*, 2:310–355, 1992.

[281] G. Tononi, O. Sporns, and G.M. Edelman. A measure for brain complexity: Relating functional segregation and integration in the nervous system. *Proc Natl Acad Sci U.S.A*, 91:5033–5037, 1994.

[282] N. Trujillo-Barreto, E. Aubert-Vazquez, and W.D. Penny. Bayesian M/EEG source reconstruction with spatio-temporal priors. *Neuroimage*, 39:318–335, 2008.

[283] N.J. Trujillo-Barreto, E. Aubert-Vazquez, and P.A. Valdes-Sosa. Bayesian model averaging in EEG/MEG imaging. *Neuroimage*, 21:1300–1319, 2004.

[284] C. Tsallis. Possible generalization of boltzmann-gibbs statistics. *Journal of Statistical Physics*, 52:479–487, 1988.

[285] M. Ullsperger and S. Debener, editors. *Simultaneous EEG and fMRI Recording, Analysis and Application*. Oxford University Press, 2010.

[286] F. Varela, J-P. Lachaux, E. Rodriguez, and J. Martinerie. The brainweb: Phase synchronization and large-scale intregration. *Nat Rev Neurosci*, 2:229–239, 2001.

[287] J.A. Vastano and H.L. Swinney. Information transport in spatiotemporal system. *Phys Rev Letters*, 60:1773–1776, 1988.

[288] B.D. Van Veen, W. Van Drongelen, M. Yuchtman, and A. Suzuki. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on Biomedical Engineering*, 44:867–880, 1997.

[289] R. Vincente, M. Wibral, M. Lindner, and G. Pipa. Transfer entropy – a model-free measure of effective connectivity for the neurosciences. *J Comput Neurosci*, 30:45–67, 2011.

[290] D.O. Walter. Spectral analysis for electroencephalograms: Mathematical determination of neurophysiological relationships from records of limited duration. *Exp Neurol*, 8:155–181, 1963.

[291] R.D. Walter and C.L. Yeager. Visual imagery and electroencephalographic changes. *Electroencephalogr Clin Neurophysiol*, 8:193–199, 1956.

[292] W.G. Walter, R. Cooper, V.J. Aldridge, W.C. McCallum, and A.L. Winter. Contingent negative variation: An electric sign of sensorimotor association and expectancy in the human brain. *Nature*, 203:380–384, 1964.

[293] L. Wang, X. Liu, K.G. Guise, R.T. Knight, J. Ghajar, and J. Fan. Effective connectivity of the fronto-parietal network during attentional control. *J Cogn Neurosci*, 22:543–553, 2009.

[294] X. Wang, Y. Chen, and M. Ding. Estimating granger causality after stimulus onset: A cautionary note. *Neuroimage*, 41:767–776, 2008.

[295] D.J. Watts and S.H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393:440–442, 1998.

[296] P.D. Welch. The use of fast fourier transform for the estimation of power spectra: A method based on time averages over short, modified periodograms. *IEEE Trans Audio and Electroacoustic*, 15:70–73, 1967.

[297] J.C. Woestenburg, M.N. Verbaten, and J.L. Slangen. The removal of the eye-movement artifact from the EEG by regression analysis in the frequency domain. *Biological Psychology*, 16:127–147, 1983.

[298] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. Vaughn. Brain-computer interface technology: A review of the first international meeting. *IEEE Trans. Rehab. Eng.*, 8:164–173, 2000.

[299] C.C. Wood, D. Cohen, B. N. Cuffin, M. Yarita, and T. Allison. Electrical sources in human somatosensory cortex: Identification by combined magnetic and potential recording. *Science*, 227:1051–1053, 1985.

[300] J. Xu, Z. Liu, R. Liu, and Q. Yang. Information transmission in human cerebral cortex. *Physica D*, 106:363–374, 1997.

[301] S.M. Zeki. Functional specialisation in the visual cortex of the rhesus monkey. *Nature*, 274:423–428, 1978.

[302] Z. Zhang and D.L. Jewett. Insidious errors in dipole localization parameters at a single time-point due to model misspecification of number of shells. *Electroencephalogr Clin Neurophysiol*, 88:1–11, 1993.