INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Statistical Analysis of Dynamic Interdependence Patterns in the Cortex

by

Wilson A. Truccolo-Filho

A Dissertation Submitted to the Faculty of the
Charles E. Schmidt College of Science
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

Florida Atlantic University
Boca Raton, Florida
December 2001

UMI Number: 3029070

UMI Microform 3029070
Copyright 2002 by Bell & Howell Information and Learning Company.
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Statistical Analysis of
Dynamic Interdependence Patterns in the Cortex

by

Wilson A. Truccolo-Filho

This dissertation was prepared under the direction of the candidate's dissertation advisors, Dr. Mingzhou Ding and Dr. Steven L. Bressler, Program in Complex Systems and Brain Sciences, and has been approved by the members of his supervisory committee. It was submitted to the faculty of The Charles E. Schmidt College of Science and was accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

SUPERVISORY COMMITTEE:

Dissertation Advisor

Dissertation Advisor

Chairperson, Program in Complex Systems and Brain Sciences

Dean, Charles E. Schmidt College of Science

Dean of Graduate Studies and Research

11/13/01

Date

Acknowledgements

In the spring (South Hemisphere) of 1995 I was fortunate to read an exciting book titled Dynamic Patterns: the Self-Organization of Brain and Behavior. That happened in a time I had already been looking at several US universities for a Ph.D. training program focused on the application of dynamical system theory to the study of brain function. I had received a doctoral fellowship from the Brazilian government and decided then to apply to the program in Complex Systems and Brain Sciences, at Florida Atlantic University.

My heartiest thanks go to Scott Kelso and Betty Tuller, faculty, staff members and students of the Center for Complex System and Brain Sciences for the invaluable opportunity of studying and working here in the past five years.

In my research work and in the preparation of this dissertation I have counted with the collaboration of several people, in special: my advisors Dr. Mingzhou Ding and Dr. Steven Bressler, members of the dissertation committee Dr. Viktor Jirsa and Dr. Dawei Dong, and research collaborators Dr. Kevin Knuth, Dr. Govindan Rangarajan and Dr. Maciek Kaminski. Also, I am thankful to the Brazilian Research Council, CNPq, for funding my first four years here and to friends in Brazil who introduced me to this field of research: Lea Fagundes, Marcus Basso, Waldemar Paim and Elton Ferlin.

My especial thanks go to Mingzhou, Steve, Dawei and Gautam, with whom I had the privilege of innumerous moving and inspiring discussions.

I dedicate this work to the ones I have missed the most: Wilson, Nelcy, Cristiane, Eliane, Juliane, William, Ricardo, Augusto, Artenio, Deroni, and in memory of Susana Klock.

ABSTRACT

Author:	Wilson A. Truccolo-Filho
Title:	Statistical Analysis of Dynamic Interdependence Patterns in the Cortex
Institution:	Florida Atlantic University
Dissertation Advisors:	Mingzhou Ding Steven L. Bressler
Degree:	Doctor of Philosophy
Year:	2001

This dissertation is an investigation of the sources of commonly observed event-related transients in statistical measures of interdependence: variance, cross-correlation, power spectrum density and coherence spectrum density time functions. These measures are often employed in the analysis of spatio-temporal interdependence patterns in neural activity.

In order to understand the phenomenon, the origins of the variability of event-related responses are revisited. The time series of single trial cortical event-related potentials typically have a random appearance, and their trial-to-trial variability is commonly explained by the classic signal-plus-noise model, in which random ongoing background noise activity is linearly combined with a stereotyped evoked response. Here,

we demonstrate that more realistic models, challenging both the linear superposition and the trial-to-trial stationarity of the event-related responses, can account for such event-related transients. In particular, two effects are considered: the nonlinear gain modulation in neural networks coupled through sigmoid functions and the trial-to-trial variability in amplitude and latency of the event phase-locked responses. An extensive analysis and characterization of both effects in interdependence measures is carried out through both analytical and numerical simulations in Chapter 2. Chapter 3 presents the outcome of testing the predicted effects on LFP data recorded from implanted intracortical electrodes in monkeys performing a visuo-motor pattern discrimination task. Overall, the results point to a large contribution of the trial-to-trial variability of event phase-locked responses on the observed event-related transient in statistical interdependence measures. Because variability of the event-related responses is commonly ignored, event-related modulations in power spectral density, cross-correlation, and spectral coherence are often attributed to dynamic changes in functional connectivity within and among neural populations. It becomes then crucial the separation or removal of the trial-to-trial amplitude and latency variability effect from the statistical measures. In order to achieve this goal, the reconstruction of the single trial event phase-locked potentials is required. In Chapter 4, we approach this problem from a Bayesian inference perspective. The posterior probability density is derived for a specified number of event phase-locked components using data from single or multiple sensors. The *Maximum A Posteriori* solution is used to obtain the phase-locked component waveforms and their single trial

parameters. The outcome is a further and definitive support for predominance of the effect of the nonstationarity of the phase-locked responses on the statistical quantities.

Based on the theoretical and experimental analysis conducted in Chapters 2,3 and 4, a framework for the statistical analysis of dynamic spatio-temporal interdependence patterns in Local Field Potential data is articulated.

Table of Contents

Chapter 1 Introduction **1**

1.1 Spatio-Temporal Interdependence Patterns in the Cortex 1

1.2 Framework for the Statistical Analysis of Dynamic Interdependence Patterns 13

1.2.1 Probabilistic Models **13**

1.2.2 Linear Stochastic Models of Ongoing Activity **19**

Chapter 2 Sources of Event-Related Transients in Interdependence Measures **38**

2.1 Nonlinear Gain Modulation Effect 38

2.2 Trial-to-trial Nonstationarity of Event-Related Phase-Locked Responses 49

2.2.1 The VSPN Model and its Predictions for the Single Channel Quantities **50**

2.2.1.1 Amplitude Variability 50

2.2.1.2 Latency Variability 52

2.2.1.3 Amplitude and Latency Variability 54

2.2.2 Implications of the VSPN Model for Statistical Interdependence Measures Involving Two Channels **55**

Chapter 3 Event-Related Transients in Interdependence Measures: Experimental Results **65**

3.1 Methods 65

3.1.1 Visuo-Motor Pattern Discrimination Task **65**

3.1.2 Recording and Data Preprocessing **66**

3.1.3 AMVAR Spectral Estimation **67**

3.1.4 Absolute Phase Histograms **68**

3.1.5 Testing for Departure from Uniform Phase Distribution **71**

3.1.6 Cross-Correlation Time Function **71**

3.2 Single Channel Quantities: AERP, Variance and Power 72

3.3 Cross-Correlation and Spectral Coherence 83

3.4 Discussion..... 92

3.4.1 Variability of Cortical Recordings **93**

3.4.2 Observed Effects of the Trial-to-Trial Variability on Recorded Local Field Potentials **95**

3.4.3 Separation of Stimulus Phase-Locked and Non-Phase-Locked Components in Cortical Recordings **97**

Chapter 4 Bayesian Analysis of Single Trial Cortical Event-Related Components . 99

4.1 Generative Model and Single Trial Parameter Estimators 100

4.2 Algorithm Implementation 112

4.3 Separation of Event-Related Components and Ongoing Activity: Identification of the Origin of Event-Related Increase in Ensemble Variance Time Function. 117

4.4 Removing the Effect of Trial-to-Trial Variability of Event Phase-Locked Responses from the Statistical Measures of Interdependence... 126

4.5 Discrimination Between Simultaneous Cortical Processes 127

4.6 Generative Model in the Frequency Domain 130

4.7 Discussion..... 133

Chapter 5 General Discussion 139

Appendices..... 144

6.1 Appendix A: Stability Conditions for Equilibrium Points in Oscillatory Neural Networks..... 145

6.1.1 Models.... 145

6.1.2 Local Stability Analysis of Model A 148

6.1.2.1 Symmetric Coupling Matrix..... 149

6.1.2.2 Non Symmetric Coupling Matrix..... 152

6.1.3 Stability Analysis of Model B 156

6.1.4 Relations to the Hopfield Neural Network Model and Global Stability 162

6.1.5 The Lienard-Chipart Criterion..... 163

6.1.6 The Generalized Routh-Hurwitz Criterion 164

6.2 Appendix B: Bifurcation Scenarios..... 165

6.2.1 Single Column..... 165

6.2.1.1 Bifurcations Under Input..... 165

6.2.1.2 Frequency and Phase 168

6.2.2 Two Coupled Columns 169

6.2.2.1 Properties of the Origin in the Absence of Input..... 170

6.2.2.2 Bifurcations Under Input..... 173

The homogeneous case: $k_{21} = k_{12}$ and $I_1 = I_2$. 173

The heterogeneous case: $k_{21} \neq k_{12}$ and/or $I_1 \neq I_2$. 177

***6.2.3 N*-Coupled Columns** **179**

6.2.3.1 Bifurcations Under Input 179

References **182**

Curriculum Vitae **197**

Table List

Chapter 3

3.1 Maximum Kuiper V during post-stimulus period: 0-200 ms.....79

Chapter 4

4.1 Correlation between single trial parameters and single trial pre-stimulus power.....125

Figure List

Chapter 1

Figure 1-1: Event-related transient modulation of spectral coherence time functions. 4

Figure 1-2: Single trial residual time series..... 6

Figure 1-3: Latent variable model for single channel cortical LFPs..... 18

Figure 1-4: Comparison between AMVAR spectral estimation and non-parametric high-resolution multitaper spectogram. 33

Chapter 2

Figure 2-1: Power and coherence spectrum density dependence on effective coupling strength K_1. 44

Figure 2-2: Transient nonlinear gain modulation of power and coherence time functions. 47

Figure 2-3: Single trial example, AERP and coherence for the same system and variables simulated in the previous Figure 2-2. 48

Figure 2-4: Simulation using a sine wave signal of variable amplitude (with independent noise set to zero). 53

Figure 2-5: Relation between ensemble average, $< z_r(t) >_r$, and ensemble variance time function, $\sigma^2(t)$, for the case of trial-to-trial latency variability in simulated time series..... 56

Figure 2-6: Relation between ensemble average, $< z_r(t) >_r$, and ensemble variance time function, $\sigma^2(t)$, for the case of trial-to-trial amplitude and latency variability in simulated time series.. 57

Figure 2-7: Time-dependent signal-to-noise ratio effect on the modulation of the cross-correlation time function: transient increase I.. 62

Figure 2-8: Time-dependent signal-to-noise ratio effect: transient increase II.. 63

Figure 2-9: Time-dependent signal-to-noise ratio effect: transient decrease.. 64

Chapter 3

Figure 3-1: Sites for the transcortical electrodes placement.. 69

Figure 3-2: Example time functions of ensemble mean (AERP), variance (σ^2 (t)) and power at 12 Hz.. 74

Figure 3-3: Time of the largest peak in the variance function plotted against the time of the largest AERP extremum.. 75

Figure 3-4: Example of an AERP function and power functions from a pre-striate (pst2) site from the monkey GE..... 76

Figure 3-5: Event-related modulation of ongoing ~ 20 Hz oscillation.. 77

Figure 3-6: Four representative Kuiper V time functions.. 80

Figure 3-7: Bimodal phase.. 81

Figure 3-8: Uniform phase. 82

Figure 3-9: Oscillatory modulation of cross-correlation time functions.. 85

Figure 3-10: Peak-time of the lagged point by point cross-correlation time function versus the peak-time of the variance function for the channel with the earliest peak in variance in the pair.. 86

Figure 3-11: Event-related changes in power and coherence spectrum time functions I.. 87

Figure 3-12: Event-related changes in power and coherence spectrum time functions II.. 88

Figure 3-13: Event-related changes in coherence spectrum time functions.. 89

Figure 3-14: Event related transient increase in 12Hz coherence in monkey GE: summary.. 90

Figure 3-15: The time-dependent signal-to-noise ratio effect on the coherence time function..... 91

Chapter 4

Figure 4-1: Potential application of the multi-channel multicomponent event-related LFP estimation.. 113

Figure 4-2: Estimation of single trial event-related components. 119

Figure 4-3: Estimated single trial event-related components.. 120

Figure 4-4: Single trial estimated event-related phase-locked components.. 121

Figure 4-5: Estimation of ongoing activity.. 122

Figure 4-6: Relations between components' latency and reaction time.. 123

Figure 4-7: Estimation of ongoing activity. 124

Figure 4-8: Removal of the trial-to-trial variability effect on the coherence time functions.. 128

Figure 4-9: Discrimination between processes related to two simultaneous events.. 129

Appendices

Figure 6-1: Model schematics 146

Figure 6-2: Single column bifurcation diagrams as function of the parameter Q_m 168

Figure 6-3: Frequency dependence on the rate constants a and b, obtained from the linear estimation for the frequency of the excitatory and inhibitory populations in the single column..... 170

Figure 6-4: Functions f and g used for determining the nature of the equilibrium point at $(x, y) = (0,0)$ for the 2-symmetrically coupled columns..... 172

Figure 6-5: The fixed point equation for the 2-symmetrically coupled columns with input $I = 0$..... 175

Figure 6-6: Bifurcation diagram for the 2 symmetrically coupled columns.. 177

Figure 6-7: Bifurcation diagram for the 2 asymmetrically coupled columns..... 178

Figure 6-8: Bifurcation diagram for the 2 asymmetrically coupled columns..... 179

Figure 6-9: Chaotic behavior of the 2 asymmetrically coupled system. 180

Figure 6-10: Types of behavior for the 64 asymmetrically coupled columns system.. ... 181

Chapter 1

Introduction

1.1 Spatio-Temporal Interdependence Patterns in the Cortex

A primary goal in neural data analysis is to describe how coordinated actions of neurons evolve in time conditioned on different behavioral events. Such coordinated action can be expressed both in terms of (spatial) interdependence patterns among neurons and in terms of (temporal) interdependence patterns in their current and past activities.

This goal is motivated in part by the fact that a number of cerebral cortical functions are thought to involve transient task-dependent changes in the functional interdependence of neurons in the same and/or different cortical areas. Examples include

perceptual binding and segmentation (Gray, 1999; von der Malsburg, 1994), multimodal integration (Bressler, 1995; Bressler, 1996; Damasio, 1989a; Damasio, 1989b; Sporns, Tononi, & Edelman, 1994; Tononi, Sporns, & Edelman, 1992), selective attention (Fries, Reynolds, Rorie, & Desimone, 2001; Olshausen, Anderson, & Van Essen, 1993), compositional representation (Bienenstock, Geman, & Potter, 1997), recognition (Hopfield & Brody, 2001), and coordination dynamics (Kelso, 1995). Also, theories of optimal coding in sensory systems invoke dynamic changes in the spatio-temporal properties of the neural activity in order to achieve optimal transmission of relevant information (Attick, 1992; Dong & Atick, 1995; Olshausen & Field, 1996; Truccolo & Dong, 2001).

Generally, evidence for such changes purportedly comes from stimulus- or task-related modulation of statistical interdependence measures, such as cross-correlation or spectral coherence, that are derived from various types of cortical recordings. In addition, changes in the level of interaction within a neuronal population are inferred from temporal modulation of the power spectral density of population signals such as the local field potential (LFP) or electroencephalogram (EEG) (Fries, Reynolds, Rorie, & Desimone, 2001; Kalcher & Pfurtscheller, 1995; Pfurtscheller & Lopes da Silva, 1999). Indeed, recent studies have provided evidence for the event-related modulation of all these statistical measures in a number of different experimental preparations. Event-related modulations in cross-correlation and spectral coherence have been usually understood in terms of event-related changes in coupling strength between neuronal components, leading subsequently to changes in the level of synchrony of neuronal pulse

activity. Fast transient changes in functional connectivity of the underlying neural circuitry (i.e. changes in synaptic efficacy and/or effective connectivity occurring on time scales of ~ 100 ms) are often pointed out as the sources of such events (Aertsen, Erb, & Palm, 1994; Aertsen, Gerstein, Habib, & Palm, 1989; Buchel & Friston, 1997; Chawla, Rees, & Friston, 1999; Gray, Konig, Engel, & Singer, 1989; Rodriguez, et al., 1999; Srinivasan, Russell, Edelman, & Tononi, 1999; Tallon-Baudry, Bertrand, Delpuech, & Pernier, 1997; Tallon-Baudry, Bertrand, Peronnet, & Pernier, 1998).

This dissertation has evolved as the result of the attempt to understand the origin of these transient modulations in this particular type of experimental paradigm: the event-related potentials recordings. Figure 1-1 shows an illustrative example of this phenomenon.

In order to understand the difficulties involved, we need to reconsider the long-standing question of how to model the cortical potentials that follow presentation of a sensory stimulus. At least since the time of Dawson (Dawson, 1954), the most common model of the post-stimulus cortical potential postulates that a recorded single trial potential is the linear combination of an invariant stimulus phase-locked evoked component (signal) and an ongoing background noise component. The evoked response component is treated as being time-invariant since any changes over trials in its amplitude, latency onset, or waveform are considered to be negligible. The usual method of measuring this component is to repeatedly present the same sensory stimulus and then average the post-stimulus potentials across an ensemble of trials (ensemble average). Averaging is presumed to decrease the size of the noise component, while leaving the

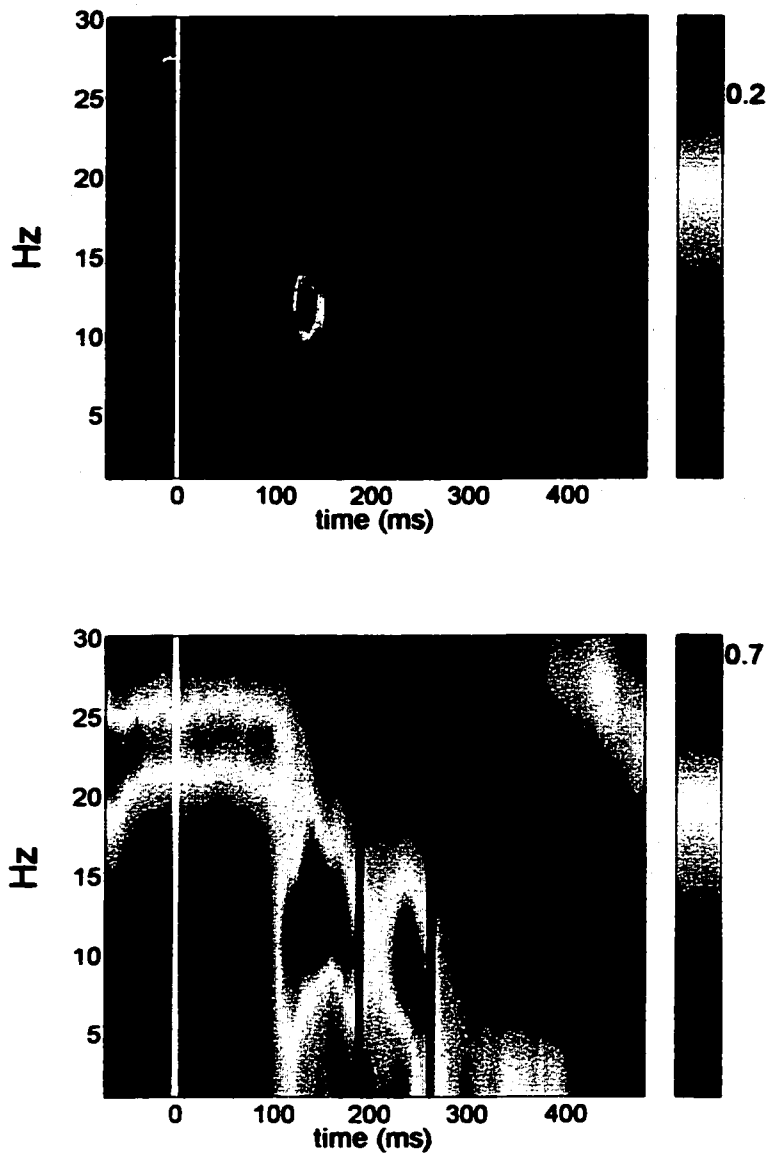

Figure 1-1: Event-related transient modulation of spectral coherence time functions. Examples from two pairs of cortical sites: striate (str3) and pre-frontal sites (top plot) and parietal sites (bottom plot). LFPs were recorded intracortically from a monkey performing a GO-NOGO visuo-motor pattern discrimination task. Stimulus onset is at time zero and coherence plots are computed for the GO-trials and a particular visual pattern. Notice that shortly after 100 ms, there is a transient increase in the 12 Hz coherence in both pairs. The parietal pair is even more interesting since there is a switch from coherence around 22 Hz to 12 Hz just after the initial sensory activation. After the event-related response is over, the 22 Hz coherence returns. The green line is the estimated mean time for the GO-NOGO decision and the red line gives the mean reaction time for the GO trials.

signal component unchanged, thereby enhancing the signal-to-noise ratio. Since averaging is time-locked to a particular event onset, the result is called the average event-related potential (AERP). From this still widely accepted conceptualization (McGillem & Aunon, 1987), henceforth referred to as the signal-plus-noise (SPN) model, it follows that: (1) all the variability in single trial recordings is due to the independent ongoing noise component; (2) the AERP asymptotically approaches the true invariant evoked response as the number of trials involved in the averaging increases; and, most importantly, (3) when the AERP is subtracted from single trial recordings, the resulting residual time series do not contain significant event-related information. Rigorous acceptance of the SPN model assumptions would imply that no event-related modulation should be observed in statistical measures that are, by definition, computed on the residual time series, i.e. on time series from which the average event-related potential (AERP) has been removed. An example of a single trial residual time series is given in Figure 1-2. These measures include cross-correlation, spectral coherence and power spectrum density time functions. However, as pointed out above, there is abundant experimental evidence that event-related modulation does occur, suggesting that the SPN model is significantly violated in its main principles: (1) the ongoing activity can be event-related as a consequence of, for example, nonlinear interactions between the evoked response and the ongoing activity, and/or (2) the evoked response may vary from trial to trial. In this dissertation we explore both cases.

First, in the case of accepting nonlinear interactions of the stimulus evoked response and ongoing activity, the ongoing component could modulate the stimulus-

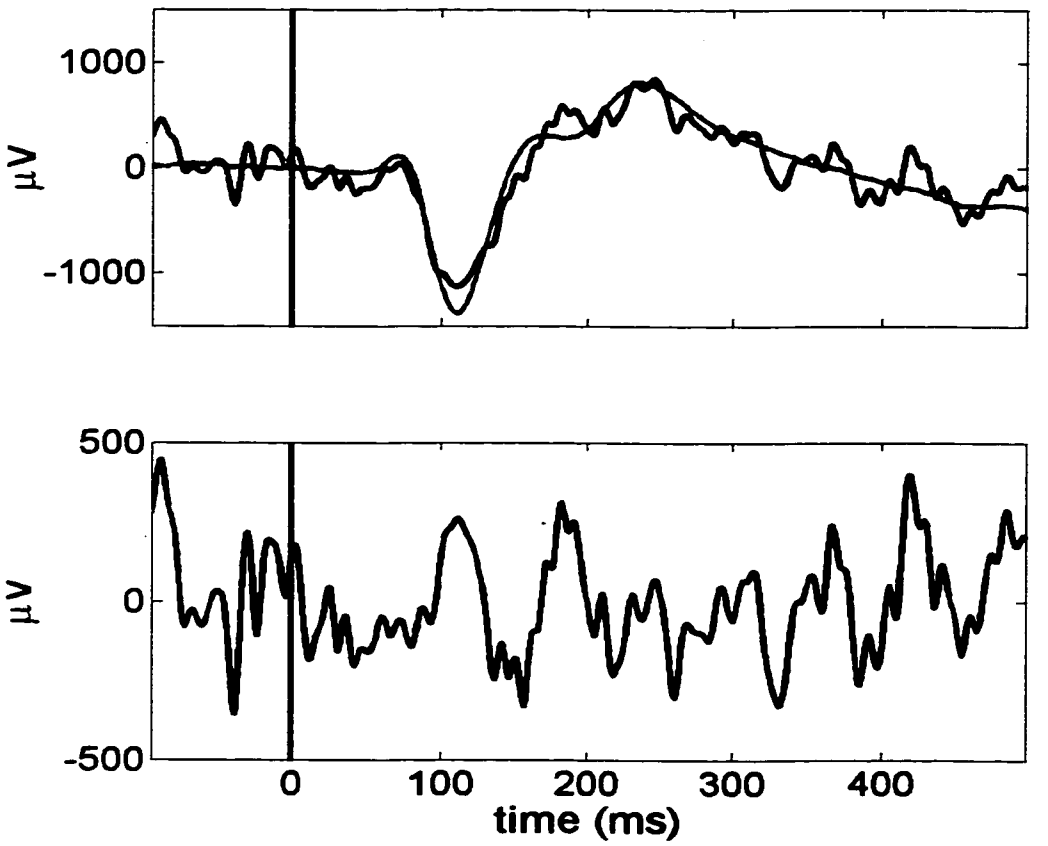

Figure 1-2: Single trial residual time series. The statistical measures of interdependence are computed on the ensemble mean subtracted time series. In the top plot, a single trial LFP time series recorded intracortically from a striate site in a monkey performing a GO-NOGO visuo-motor pattern discrimination task. The ensemble average or AERP is shown in red. In the bottom plot, the resulting residual time series after the AERP subtraction. Stimulus onset is at time zero.

evoked response, and vice versa. A simple illustration is provided in the following scenario. Dynamical models of neuronal population activity relate the field potential and pulse density (the number of spikes per unit volume) by a nonlinear sigmoid function (Eeckman & Freeman, 1991a; Freeman, 1975). A resulting property of networks of such populations is the dynamic modulation of gain and effective connectivity by the network's mean activity level. This modulation can affect both the local population properties and the interactions among local populations. Transient changes in the population's mean level of activity, like that produced by stimulus-evoked responses, can change the gain of the population's output sigmoid function, resulting in amplification or attenuation of its ongoing activity. This effect can be measured by transient changes in the time course of the ensemble variance time function and in power spectrum density time functions of the residual time series. Such changes are likely to be observed over a range of frequencies, especially if the pre-stimulus activity is broadband. Also, as the gain level of the output sigmoid function of a population changes, its effective connectivity with other populations may also be modulated, leading to possible changes in the level of synchrony between the neuronal population activities. In this way, the residual time series from two interacting populations can exhibit event-related modulation of their statistical interdependence, as measured by fast transients in the time course of cross-correlation and coherence time functions. This mechanism will henceforth be referred to as the nonlinear gain modulation effect. Other nonlinear mechanisms, based on the effect of the background level of activity and evoked transients on effective connectivity (Aertsen, Erb, & Palm, 1994; Aertsen, Gerstein, Habib, & Palm, 1989; Chawla, Lumer, & Friston,

1999; Chawla, Lumer, & Friston, 2000; Friston, 2000) are similar in spirit. Besides being a natural consequence in models of sigmoidally coupled neuronal populations, the effect has potential functional relevance: slow spatio-temporal patterns of activity (say in the alpha range) could determine the mean level of activity and subsequently soft-wire the connectivity between neural components. The result would be a mechanism for dynamic functional connectivity at either broad or specific frequency bands (say gamma band). Changes in intrinsic parameter regulating the frequency characteristics of neuronal populations could also lead either to resonance or to decoupling between the populations (Baird & Eeckman, 1993; Hoppensteadt & Izhikevich, 1999; Izhikevich, 1999), producing similar event-related modulations in interdependence measures.

Second, a concurrent, but not exclusive, explanation resorts to the possibility of nonstationarity of the event-related responses and its consequences. Even though stationarity is usually assumed as a pre-condition, we argue that in practice it might be unattainable in most of the experimental designs involving behaving animals. Temporal modulation of power and interdependence measures may result from *trial-to-trial* nonstationarity of the cortical evoked response. Trial-to-trial variability in amplitude, latency and waveform of the evoked response constitute manifestations of this nonstationarity (Coppola, Tabor, & Buchsbaum, 1978; Lange, Pratt, & Inbar, 1997; Mocks, Gasser, Tuan, & Kohler, 1987; Pham, Mocks, Kohler, & Gasser, 1987). We will demonstrate that trial-to-trial variability in the amplitude and latency of evoked local field potentials, can, in the appropriate context, lead to temporal modulation of the aforementioned statistical measures, producing effects that could resemble either transient

synchronization or, alternatively, transient desynchronization events. Consequently, inter-trial variability of the evoked response may appear as *intra-trial* stimulus- or task-related modulation of intrinsic parameters in the neural system. To overlook this possibility, as is commonly done, may result in the erroneous interpretation of trial-to-trial nonstationarity as *intra-trial* task-related changes in functional connectivity. For the purpose of this work we focus on the consequences of relaxing the SPN model's assumption of invariant amplitude and latency of the phase-locked event-related responses over trials. The alternative model, henceforth referred to as the Variable Signal Plus ongoing Noise activity (VSPN) model, asserts that the stimulus-triggered response has a stereotyped waveform with variable amplitude and latency onset across trials. According to the VSPN model, after subtracting out the AERP, the single trial residual time series contain two components: (1) a stimulus phase-locked component resulting from the trial-to-trial amplitude and latency variability; and (2) an ongoing noise component. Because the VSPN model includes a residual stimulus phase-locked component, it differs from the SPN model in its prediction concerning the post-stimulus modulation of power spectral density and statistical interdependence time functions. In the SPN model, the residuals remaining after AERP subtraction consist only of event-independent noise. Therefore, it predicts that the ensemble variance will be constant for all post-stimulus times. By contrast, the VSPN model predicts that the ensemble variance time function will of necessity be nonstationary, with its time course modulated according to the AERP waveform. Since the AERP is often oscillatory with clear characteristic frequencies, the phase-locked component remaining in the residuals will contribute to the power spectral

density time function at these frequencies. Consequently, the power spectral density of the LFP or EEG time series, after subtraction of the AERP, and computed as a function of time in a sliding short window, can nonetheless be modulated according to the AERP waveform. Moreover, the ratio of the variance of the residual stimulus phase-locked component to the variance of the ongoing component may change in time according to the AERP profile, especially if the time dependence of the variance of the ongoing component is weak. This means that if the residual time series from two recording channels co-vary, the time function of measures of those channels' statistical interdependence will also be modulated accordingly to the AERP profile. We refer to this type of modulatory effect on any such statistical measure as the *time-dependent signal-to-noise ratio effect*, where the signal refers again to remnants of the stimulus phase-locked components in the residual time series. The existence of phase-locked components in the residual leads to a second contrast between the two models, which will be important to assert the source of the variability in the event-related phase locked components. Specifically, the VSPN model predicts that peaks in the single trial stimulus-locked component will be either larger or smaller than the average, resulting, after AERP subtraction, in the peaks of the phase-locked component in the single trial residual time series having either positive or negative polarity. If the latency variability is negligible, then the phases of the Fourier components, at frequencies corresponding to the main characteristic frequencies of the AERP, are predicted to have a bimodal distribution. This again is in contrast with the SPN model, which predicts a flat phase distribution since the residuals are considered to originate only from the independent ongoing activity.

Given the postulation of these two possible explanations of the origins of fast transient event-related modulation of interdependence measures, the objectives of this dissertation are threefold:

(1) Provide a detailed analysis and characterization of the effects of nonlinear gain modulation effect (in models of sigmoidally coupled neurons) and of the trial-to-trial variability of the event phase-locked responses on measures of statistical interdependence: variance, cross-correlation, power and coherence spectrum density time functions.

(2) Test the existence of the derived effects in (1) on a local field potential data set.

(3) Provide a statistical framework able to isolate the contribution of the trial-to-trial variability of the phase-locked components to the statistical measures of interdependences.

To address the three objectives above, the dissertation is organized as follows. A detailed analysis and characterization of the nonlinear gain modulation and trial-to-trial variability is carried out through both analytical and numerical simulations in Chapter 2. The effects from each factor on the single channel quantities, variance and power time functions, and for the phase distribution are derived first. Then, the predictions for time functions of statistical interdependence measures (cross-correlation and coherence) are

presented, with the nonlinear gain and signal-to-noise ratio effects illustrated using simulated time series. Chapter 3 presents the outcome of testing these predictions on LFP data recorded from implanted intracortical electrodes in monkeys performing a visuo-motor pattern discrimination task (Bressler et al. 1993). Overall, the results point to a large contribution of the signal-to-noise ratio effect on the observed event-related transient in statistical interdependence measures. It becomes then crucial the separation or removal of the trial-to-trial amplitude and latency variability effect from the statistical measures. In order to achieve this goal, the reconstruction of the single trial event-related phase locked potentials is considered. If one is able to reconstruct such potentials, then the ensemble average of the aforementioned statistical measures, when calculated on the new residual time series after removing the event-related phase-locked potentials on a trial by trial basis, will exhibit no time behavior that is characteristically related to the trial-to-trial nonstationarity of the responses. In Chapter 4 and based on the VSPN model, we apply a Bayesian inference procedure to estimate the single trial evoked responses in the LFP data set, which gives further and definitive support for predominance of the signal-to-noise ratio effect on the statistical quantities. The single trial based procedure again points to the importance of identifying the sources of nonstationary variance in LFP data. The discussion and general conclusions are presented in Chapter 5. For the sake of completeness, a thorough stability and bifurcation analysis of the neuronal population model used in the investigation of the nonlinear gain effect is given in the appendix section. The stability analysis is in itself of interest since it provides original and

significant results on the coupling constraints for the stability of both symmetrically and asymmetrically coupled neural networks of arbitrary size.

Based on the results to be further detailed in the next Chapters, we conclude this chapter with a brief exposition of an initial framework for the analysis of dynamic spatio-temporal interdependence patterns in local field potential data.

1.2 Framework for the Statistical Analysis of Dynamic Interdependence Patterns

1.2.1 Probabilistic Models

Ideally, we would like to obtain the joint probability of relevant brain states and how it evolves in time conditioned on specific events. The notion of state variables for a stochastic system can intuitively be expressed as the set of variables carrying the minimal amount of information about the past history which is sufficient, given transition rules, to uniquely determine their joint probability at future times, conditioned again on specified events. Although the terms random variables and stochastic process are going to be used, the use of "probabilities" has its epistemological ground on the Bayesian approach, where probability represents a degree of belief that a proposition is true (Jaynes, 1994).

Given the difficulties in this task, an initial and more feasible goal is to start with the description of the conditional joint probability:

$$p(\Psi(t), \Psi(t-1), \ldots, \Psi(t-T) | Event), \tag{1.1}$$

where $\Psi(t)=[\psi_1(t),\psi_2(t),\ldots,\psi_M(t)]^T$ represents the relevant brain states at time[1] t, measured from a set of T time samples $\{\Psi(t),\Psi(t-1),\ldots,\Psi(t-T-1)\}$ relative to the onset of the specified E*vent*, over discrete space samples $\{m=1,2,\ldots,M\}$. All the stationary spatio-temporal structure, i.e., dynamics, will be captured by this probability. For comparison, in the case that there is no dynamics at all, the joint conditional probability should simply factorize:

$$p(\Psi(t),\Psi(t-1),\ldots,\Psi(t-T)|\,Event)=\prod_{m=1,p=0}^{MT}P(\psi_m(t-p)|\,Event). \qquad (1.2)$$

The probability in Equation (1.1) is commonly seen in neuroscience in the context of neural coding, e.g., as a description of how the occurrence of a stimulus or behavioral action is expressed through neural activity. The reciprocal probability

$$p(Event|\Psi(t),\Psi(t-1),\ldots,\Psi(t-T)), \qquad (1.3)$$

usually studied in the context of neural decoding, is related, through the Bayesian theorem, to the probability in Equation (1.1),

$$p(\Psi(t),\Psi(t-1),\ldots,\Psi(t-T)|\,Event)\,p(Event)= \qquad .(1.4)$$
$$p(Event|\Psi(t),\Psi(t-1),\ldots,\Psi(t-T))\,p(\Psi(t),\Psi(t-1),\ldots,\Psi(t-T))$$

Notice that the above is just the joint probability:

$$p(\Psi(t),\Psi(t-1),\ldots,\Psi(t-T),Event). \qquad (1.5)$$

Following the discussion in the previous section, it is now evident that there are two main challenges in the analysis of spatio-temporal interdependence patterns: a) latent

---

[1] The notation for discrete or continuous time should be clear from the context.

variable issues stemming from the fact that our measures commonly record mixed signals originated from many different processes, e.g., sensory activations, integration, motor execution, and other unrelated ongoing processes; and b) the dual nonstationarity of the brain processes; on the one hand, even a simple pattern discrimination task involves many different stages, including preparation, sensory activation, decision and execution, that can all happen in very short time intervals of less than 350 ms; on the other hand, in the experimental paradigm of event-related responses we face yet another type of nonstationarity stemming from the trial-to-trial variability of the event-related responses themselves.

The problem of latent variables arises from the fact that we do not measure directly the relevant brain states $\Psi(t)$. Instead our measurement vector $Z_r(t)$ results from a mixture of signals related to brain states, measurement noise and other unpredictable components. When approaching this problem, a first step is to start with a measurement or generative model specific to the experimental paradigm.

For the case of event-related potential measurements, three main components have for a long time been taken as the main contributors to the recordings: the stimulus or event phase-locked signals, the uncorrelated spontaneous activity and induced activities. The most desirable approach would be to account for all these components in a unified way, i.e., as the result of the behavior of dynamical models of neuronal populations. Initial attempts in this direction were taken by Freeman (Freeman, 1975), who modeled the phase-locked signals (estimated by the AERP) as resulting from the impulsive behavior of specific dynamical systems. There are however difficulties in following here

this approach. For example, the type of data will be employed in this dissertation consists of very localized local field potentials. Yet the possibility of having recorded time series containing mixed activities from neuronal populations with different dynamical properties is still very likely. Also, experimental data that would allow for a detailed parametric modeling of neuronal activity are basically unavailable. In the absence of more detailed information about the dynamical properties of the neuronal population generating the time series, a simpler data driven approach is adopted at this stage. A generative model developed in this way should capture at least four main properties of LFP recordings in the event-related paradigm: a) the existence of signals relatively phase-locked to a specific event onset; b) the trial-to-trial variability of amplitude and latency of the event-related signals or components; c) the possibility that event-related responses may contain multi components with differential variability of their amplitudes and latencies; and d) the existence of event unrelated signals (spontaneous activities plus measurement noise) and also non-phase locked event-related or induced spontaneous activities.

For the case of a single channel recording (see Chapter 4 for the full multivariate treatment), the following generative model is employed:

$$z_r(t) = \sum_{n=1}^{N} a_{nr} s_n(t - \tau_{nr}) + \eta_r(t), \qquad (1.6)$$

where $z_r(t)$ is the LFP recording in the r^{th} trial, $s_n(t)$ is the n^{th} event-related component waveform with a trial-to-trial variable amplitude and latency given, respectively, by a_{nr} and τ_{nr}. The unpredictable component $\eta_r(t)$, henceforth simply referred as ongoing

activity, denotes a zero mean activity including spontaneous activities and other uncertainties such as measurement noise. A more detailed treatment of the modeling of the ongoing activity is provided in the next section.

A graphical representation of this generative or latent variable model is shown in Figure 1-3. The waveforms $s_n(t)$ and their correspondent amplitudes a_n and latencies τ_n have their probability distributions specified by a set of parameters $\theta(t)$. Similarly, the ongoing activity $\eta(t)$ has its distribution specified by a set of parameters $\theta_\eta(t)$. These parameters can be a function of time: that is the case where, for example, the ongoing activity is nonstationary. All those variables and parameters are latent or hidden, since the only direct observable is the measurement itself: $z(t)$. The rectangles containing the variables represent the sampling, which generates each sample trial r and sample time point t. The self-loop in the ongoing activity represents time dependence between successive time points.

The analysis framework proposed here thus involves two main aspects. First, a prior task in the statistical analysis of interdependence patterns is then to separate these signal components and ongoing activities. It is fundamental to understand that the problem *is not the one of mathematical inversion, but of statistical inference, i.e, probability calculus and inductive logic.* The details of the inferential approach as well as its relations to other decomposition techniques like PCA, ICA and wavelet like filters are shown in the Chapter 4.

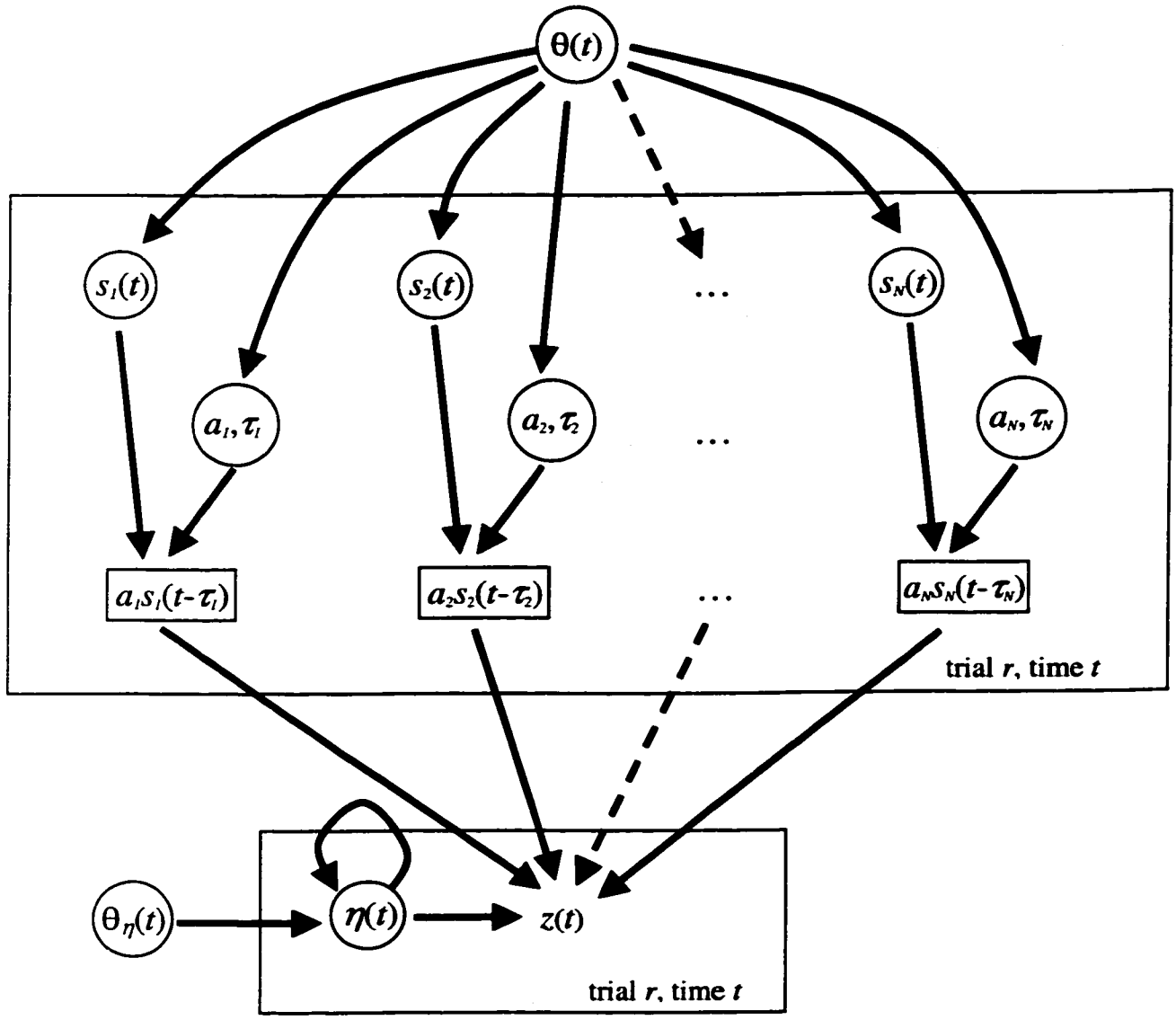

Figure 1-3: Latent variable model for single channel cortical LFPs. The generative model for event-related LFP measurements $z(t)$ consists of a linear combination of latent variables, namely event-related phase-locked components $s_n(t)$ and ongoing activities $\eta(t)$. The event-related components have trial-to-trial variable amplitudes a_n and latencies τ_n. All the latent variables are assumed to be generated from a particular distribution specified by the parameters $\boldsymbol{\theta}(t)$ and $\boldsymbol{\theta}_\eta(t)$. The rectangles represent the sampling process at trial r and time t.

Second, once the measured time series are decomposed, the analysis of spatio-temporal interdependence patterns should be based on the models of the phase-locked event-related components and on the analysis of the ongoing activities. Interdependences among the phase-locked components should be investigated by correlation analysis on their trial-to-trial amplitudes and latencies. Interdependence analysis of the ongoing activity is a more complex issue that is addressed in the next section.

1.2.2 Linear Stochastic Models of Ongoing Activity

The interdependence analysis of the estimated ongoing activity leads to additional problems. To start with, the spectral domain is preferred for the analysis of 2^{nd} order statistical structure, instead of time domain measures like auto-correlation and correlation functions. Spectral measures facilitate the discrimination between interdependences happening at different time scales and also local error bars are possible, since for stationary processes the correlations are independent in the frequency space. However, the spectral estimation presents difficulties of its own. Direct spectral estimation is known to be seriously biased (Percival & Walden, 1993), and this is particularly the case when the data segments are short. In the signal processing community, less biased estimates are obtained with the use of data windows followed by smoothing of the spectral estimator. The resulting estimator is still inconsistent and does not provide a theoretical foundation for the use of the data windows and their choice. Improvement in this direction was made with the introduction of the multitaper estimators, providing both the theoretical base for the use of tapers (windows) and also the optimal taper functions (Percival & Walden,

1993; Thomson, 1982), making this approach the current choice for the serious non-parametric spectral estimation. However, this technique is also not adequate for our purpose. Given the goal of tracking the nonstationarity of cortical processes, very short time windows need to be used. At this point, the multitaper technique is not appropriate (e.g., for a 50 ms window, frequencies under 20 Hz would be left out and the benefits of the multitaper method would be reduced since in this case a single taper function would be used). Spectral estimation based on wavelets, now common for time-frequency analysis, is also problematic since their statistical properties are not completely understood yet. More recently, a better alternative appeared with the multitaper high-resolution spectrograms (Thomson, 2000). Here, we chose instead a parametric estimation that has been developed for the analysis of short time window LFPs by Ding et al. (Ding, Bressler, Yang, & Liang, 2000). We will provide later an example showing that it provides better temporal resolution than the multitaper high-resolution, while capturing basically the same spectral structure.

Given its unpredictable nature, the ongoing activity will be modeled as a stochastic process, specifically as a discrete time continuous state Markov or autoregressive process. The probabilistic modeling will be based on the Wold decomposition theorem (Wold, 1938) and the exposition of the rationale in this approach will follow the derivations given by (Box, Jenkins, & Reinsel, 1994). The basic theorem states that any wide sense stationary (WSS) stochastic process can be decomposed into a component that is completely random and one that is deterministic. A deterministic random process is one that is perfectly predictable based on infinite past. This is to say

that $z(t)$ may be expressed as $\sum_{p=1}^{\infty} a_p z(t-p)$ without error for some constants a_p. An example of this would be the decomposition of a pure sinusoid (randomly phased to ensure WWS) and white noise into a purely random component (the white noise) and a deterministic component (the sinusoid). Alternatively, we may view this decomposition in the spectral domain as one that separates the power spectrum density into a continuous (i.e., smooth) component representing the white noise and a discrete component (impulsive behavior) representing the sinusoid (Kay, 1988).

More importantly, the theorem also states that any purely non-deterministic zero-mean stationary process $z(t)$ possesses a linear representation in the form:

$$z(t) = \varepsilon(t) + b_1\varepsilon(t-1) + b_2\varepsilon(t-2) + \ldots = \varepsilon(t) + \sum_{p=1}^{\infty} b_p \varepsilon(t-p), \qquad (1.7)$$

with $\sum_{p=1}^{\infty} [b_p]^2 < \infty$, and $\varepsilon(t)$ are uncorrelated with common variance σ_ε^2 but *need not to be independent*. This process can also be represented, under suitable conditions, as an autoregressive process, i.e. a weighted sum of past values, plus an added perturbation $\varepsilon(t)$:

$$z(t) = a_1 z(t-1) + a_2 z(t-2) + \ldots + \varepsilon(t) = \sum_{p=1}^{\infty} a_p z(t-p) + \varepsilon(t). \qquad (1.8)$$

Notice that by writing the above in the form $\varepsilon(t) = \sum_{p=0}^{\infty} a_p z(t-p)$, $a_0 = 1$, the process can be related to a filter of infinite order whose output results to be white, in other words, the filter whitens the temporal series $\{z(t), z(t-1), z(t-2), z(t-3), \ldots\}$.

The coefficients a and b may be related through a backward shift operator, defined as

$$Bz(t) = z(t-1), \quad B^k z(t) = z(t-k). \tag{1.9}$$

For example, consider the model:

$$z(t) = \varepsilon(t) - \lambda z(t-1) = (1-\lambda B)\varepsilon(t), \tag{1.10}$$

with $b_1 = -\lambda$ and $b_p = 0$ for $p > 1$. Expressing $\varepsilon(t)$ in terms of $z(t)$ we obtain

$$\varepsilon(t) = (1-\lambda B)^{-1} z(t). \tag{1.11}$$

For $|\lambda| < 1$, and after the Taylor expansion of $(1-\lambda B)^{-1}$, the above can be written as

$$\varepsilon(t) = (1+\lambda B + \lambda^2 B^2 + \lambda^3 B^3 + \ldots) z(t), \tag{1.12}$$

or in the autoregressive form:

$$z(t) = \lambda z(t-1) - \lambda^2 z(t-2) - \lambda^3 z(t-3) - \ldots + \varepsilon(t), \tag{1.13}$$

with the previous a_p expressed as $a_p = -\lambda^p$.

In general, Equation (1.7) can be written as:

$$z(t) = (1 + \sum_{p=1}^{\infty} b_p B^p)\varepsilon(t) = (\sum_{p=0}^{\infty} b_p B^p)\varepsilon(t) = b(B)\varepsilon(t), \tag{1.14}$$

with $b_0 = 1$. Similarly,

$$\varepsilon(t) = (1 - \sum_{p=1}^{\infty} a_p B^p) z(t) = a(B) z(t). \tag{1.15}$$

Finally, the relation between the coefficients a and b is obtained by operating both sides of Equation (1.15) by $b(B)$:

$$b(B)\varepsilon(t) = b(B)a(B)z(t) = z(t), \tag{1.16}$$

or

$$a(B) = b^{-1}(B). \tag{1.17}$$

Two constraints are need for the application of the representations given by equations (1.7) and (1.8): stationarity and invertibility. Stationarity is guarantee by the condition $\sum_{p=0}^{\infty} |b| < \infty$, in other words, the operator $b(B)$ must converge on or within the unit circle for $|B| \leq 1$. Invertibility is required when we are interested in establishing the dependences between present events and past happenings. Thus, we would like to move from the representation $z(t) = \varepsilon(t) + \sum_{p=1}^{\infty} b_p \varepsilon(t-p)$ to the representation in the regressive form $z(t) = \sum_{p=1}^{\infty} a_p z(t-p) + \varepsilon(t)$. For illustration, consider again the model $z(t) = (1-\lambda B)\varepsilon(t)$. Expressing the white noise as a function of the present and past of $z(t)$, we obtain

$$\varepsilon(t) = (1-\lambda B)^{-1} z_r(t) = (1+\lambda B+\lambda^2 B^2+\ldots+\lambda^p B^p)(1-\lambda^{p+1} B^{p+1})^{-1} z(t). \tag{1.18}$$

This corresponds to:

$$z_r(t) = -\lambda z(t-1) - \lambda^2 z(t-1) - \ldots - \lambda^p z(t-p) + \varepsilon(t) - \lambda^{p+1}\varepsilon(t-p-1). \tag{1.19}$$

Notice that we did not obtain yet the desired autoregressive form given in Equation (1.8) since there is an extra term $\lambda^{p+1}\varepsilon(t-p-1)$ in the expression above. However, if $|\lambda| < 1$, on letting p tend to infinity, this extra term vanishes, and we finally obtain:

$$z(t) = -\lambda z(t-1) - \lambda^2 z(t-1) - \ldots + \varepsilon(t), \tag{1.20}$$

$$z(t) = \sum_{p=1}^{\infty} a_p z(t-1) + \varepsilon(t), \tag{1.21}$$

where again $a_p = -\lambda^p$. The stationarity and invertibility conditions are independent: whatever the value of λ, Equation (1.7) specifies a stationary process. Thus, if $|\lambda| < 1$, the series

$$a(B) = (1 - \lambda B)^{-1} = \sum_{p=0}^{\infty} \lambda^p B^p \tag{1.22}$$

converges on or within the unit circle for all $|B| \leq 1$.

In general, the process in Equation (1.7) is invertible and can be represented in a autoregressive form $a(B)z(t) = \varepsilon(t)$, where $a(B) = b^{-1}(B) = 1 - \sum_{p=1}^{\infty} a_p B^p$, if

$$\sum_{p=0}^{\infty} |\lambda|^p \equiv \sum_{p=0}^{\infty} |a_p| < \infty. \tag{1.23}$$

So far, we have seen that any purely non-deterministic stochastic process can be represented in the form (1.7), and that given appropriate conditions, it can be further re-expressed in the autoregressive form (1.8). However, both representations involve infinite sets of coefficients and are in practice of little use. The goal is thus to achieve an autoregressive representation with a finite set of coefficients $\{a_p\}$, for $p = 1,2,\ldots,P$, for P small. Basically, a low order model should be found that captures practically all the spatio (for multivariate) and temporal statistical structure. Ultimately, the definite test is to check if the residual errors of the model are completely independent, i.e., the model has captured all the significant spatio-temporal statistical structure or dynamics. If the stochastic input to the model is assumed to be Gaussian, checking if the residual errors are uncorrelated would be sufficient. Another way of expressing this condition is in terms

of filter output. From a finite set of the AR coefficients, a filter is derived. If this filter when applied on the original time series results in a output time series that are completely independent in time and space (across channels), a finite order representation has been found that captures the statistical structure. In practice, parsimonic representations can be achieved in many cases by application of information theoretic criteria (Rissanen, 1989), like the Minimum Description Length, Akaike Information Criteria, and others. These criteria are later explicitly stated. Exceptions are found when dealing with long-memory processes. Those are processes where the correlations decay to zero at a slower rate than in Markov processes (Beran, 1994). In this case, the correct representation would require an infinite order AR model. Alternative approaches would include the Auto Regressive Integrated Moving Averages (ARIMA) (Box, Jenkins, & Reinsel, 1994). In practice, however, the issue is more complex and a clear-cut assertion that an underlying process has the property of long-memory is commonly not the case. When dealing with finite data sets, even low order AR models, e.g. order 5, can produce time series that will lead to signatures of long-memory processes, like approximately $1/f$ power spectrum estimates (Rangarajan & Ding, 2000).

A further restriction will be introduced for the approach undertaken in this dissertation. Notice that in the model (1.7), $\varepsilon(t)$ are uncorrelated with common variance σ_ε^2 but *need not to be independent*. This means that $\varepsilon(t)$ could refer to nongaussian processes with high order correlations. Here, $\varepsilon(t)$ will be restricted to Gaussian processes and the analysis of spatio-temporal interdependences is then limited to interdependences revealed at the level of 2nd order correlations.

In this way the ongoing activity $\boldsymbol{\eta}(t)=[\eta_1(t),\eta_2(t),\ldots,\eta_M(t)]^T$ is modeled as a Multivariate AutoRegressive (MAR) process of order P:

$$\boldsymbol{\eta}(t)=\mathbf{A}_1\boldsymbol{\eta}(t-1)+\mathbf{A}_2\boldsymbol{\eta}(t-2)+\ldots+\mathbf{A}_P\boldsymbol{\eta}(t-P)+\mathbf{E}(t), \tag{1.24}$$

$\mathbf{A}_p$ being the matrix containing the autoregressive coefficients, $A_{p,ij}=a_{p,ij}$, for $i,j=1,2,\ldots,M$. $\mathbf{E}(t)=[\varepsilon_1(t),\varepsilon_2(t),\ldots,\varepsilon_M(t)]^T$ is a multivariate zero-mean identically distributed white gaussian process with diagonal covariance matrix $\mathbf{V}$, with $V_{mm}=\sigma_m^2$.

As a side note, it is interesting to make explicit the relation of Equation (1.24) to first order Markov processes and transition probabilities. Consider the Markov process expressed as:

$$\mathbf{X}(t+1)=\mathbf{F}\mathbf{X}(t)+\mathbf{U}(t), \tag{1.25}$$

where $\mathbf{X}(t)=[\boldsymbol{\eta}(t-1),\boldsymbol{\eta}(t-2),\ldots,\boldsymbol{\eta}(t-P)]^T$, $\mathbf{U}(t)=[\mathbf{E}(t),0,\ldots,0]^T$, and the state transition matrix $\mathbf{F}$:

$$\mathbf{F}=\begin{pmatrix}\mathbf{A}_1 & \cdots & \mathbf{A}_{P-1} & \mathbf{A}_P\\ \mathbf{I} & \cdots & 0 & 0\\ \vdots & \ddots & \vdots & \vdots\\ 0 & \cdots & \mathbf{I} & 0\end{pmatrix}. \tag{1.26}$$

Also, in this case the transition probability density is simply:

$$p(\mathbf{X}(t)|\mathbf{X}(t-1))=\frac{p(\mathbf{X}(t),\mathbf{X}(t-1))}{p(\mathbf{X}(t-1))}. \tag{1.27}$$

For Gaussian processes, the above probability can be completely specified from the covariance matrices, which are derived from the coefficient matrices in $\mathbf{F}$.

The specification of the coefficient matrices and noise covariance matrix can be obtained by solving the Yule-Walker equation (Kay, 1988). A more efficient approach

adopted here is the normalized form of the Levinson, Wiggins, Robinson (LWR) algorithm (Haykin & Kesler, 1983; Morf, Vieira, Lee, & Kailath, 1978). The LWR algorithm is a multivariate extension of the maximum entropy method proposed by Burg (Burg, 1975). The intuitive rationale is as follows. When dealing with finite data sets, we always have a correlation function that is truncated at a certain time lag. In other words, the correlation function is not known after a certain time lag. The effects of this truncation are particular severe in the spectral estimation of short time data. The characteristic smearing of spectral density estimates due to this truncation can be alleviate by an explicit extrapolation of the known segment of the correlation function to the unknown segments, given certain constraints. Burg chose the maximum entropy principle (Jaynes, 1957b; Jaynes, 1957a; Sivia, 1996) as a constraint. More specifically, among the infinite possible choices of extrapolations that produce valid correlation functions, the extrapolation that should be chosen is the one that results in the maximum entropy of the time series characterized by the truncated correlation function.

Finding the adequate AMVAR parsimonic representation of a stochastic process is a problem of model selection. Several criteria for the choice of model order are given below (Haykin & Kesler, 1983; Rissanen, 1989). For the Akaike Information Criterion (AIC):

$$AIC(O) = 2\ln[\det(\mathbf{V})] + 2M^2 O / T_{total}, \tag{1.28}$$

where T_{total} is the total number of time samples of process for all the collected trials. The optimal model order is then

$$\hat{P} = \arg\min AIC(O). \tag{1.29}$$

When applying the maximum entropy estimation, the above criterion is given by:

$$AIC(O) = \ln P_M + 2M^2 O / T_{total}, \tag{1.30}$$

where P_M is the output error power of the resulting MVAR filter of order O. In the implementation used in this dissertation, we employed the Final Prediction Error (FPE) criterion:

$$FPE(O) = \frac{T_{total} + MO + 1}{T_{total} - MO - 1} P_M, \tag{1.31}$$

where the optimal model order is

$$\hat{P} = \arg\min FPE(O). \tag{1.32}$$

Another interesting criteria is the Minimum Description Length (MDL):

$$MDL(O) = T_{total} M \ln P_M + O \ln[MT_{total}], \tag{1.33}$$

with

$$\hat{P} = \arg\min MDL(O). \tag{1.34}$$

The above three criteria are related and can all be derived based on Bayesian inference and on the Shannon coding theorem (for details see Chapter 4 and (Rissanen, 1989)). A comparison between AIC and MDL is given in Rissanen, chapters 3 and 6 (Rissanen, 1989).

Once the coefficient matrices and the noise covariance matrix are known, a more useful representation of the process in Equation (1.24) is given by its spectral representation. By taking the Z-transform of Equation (1.24), we obtain:

$$\boldsymbol{\eta}(f) = \mathbf{H}(f)\mathbf{E}(f), \tag{1.35}$$

where $\mathbf{H}(f)$ is the system's transfer function:

$$\mathbf{H}(f)=\left(-\sum_{p=0}^{P}\mathbf{A}_p e^{-i2\pi pf}\right)^{-1}, \tag{1.36}$$

with $\mathbf{A}_0 = -\mathbf{I}$. The spectral density matrix results from the average:

$$\mathbf{S}(f) = <\boldsymbol{\eta}(f)\boldsymbol{\eta}^*(f)> = <\mathbf{H}(f)\mathbf{E}(f)\mathbf{E}^*(f)\mathbf{H}^*(f)>, \tag{1.37}$$

$$\mathbf{S}(f) = \mathbf{H}(f)<\mathbf{E}(f)\mathbf{E}^*(f)>\mathbf{H}^*(f), \tag{1.38}$$

$$\mathbf{S}(f) = \mathbf{H}(f)\mathbf{V}\mathbf{H}^*(f). \tag{1.39}$$

From the spectral matrix, the power spectrum density and spectral coherences (ordinary, partial and multiple) are derived. In particular the power for the m^{th} channel, and ordinary coherences between the m^{th} and the n^{th} channels, are given respectively by:

$$\Gamma_m(f) = S_{mm}(f), \tag{1.40}$$

and

$$C_{mn}(f) = \frac{|S_{mn}(f)|^2}{S_{mm}(f)S_{nn}(f)}. \tag{1.41}$$

Another important quantity is the Directed Transfer Function (DTF), expressed as:

$$D_{mn}(f) = |H_{mn}(f)|^2. \tag{1.42}$$

The DTF is not a symmetric measure and thus provides an indication of the level of statistical influence from channel n to channel m. It is a multivariate spectral representation equivalent to the bivariate time domain Granger causality (Granger, 1969; Kaminski, Ding, Truccolo, & Bressler, 2001). A normalized form is also easily obtained. A very similar quantity called the partial DTF has been also proposed by Baccala et al (Baccala & Sameshima, 2001).

The above measures reveal the second order linear statistical structure: the temporal correlations given by the power spectrum density; and the interdependence between the channel pairs by the coherences and the DTF. As mentioned before, detection of higher order temporal correlations are possible if we allow nongaussian $\varepsilon(t)$ or nonlinearities. In this case, higher order spectra, e.g. bispectrum and bi-coherence, are made possible (Nikias & Petropulu, 1993).

When dealing with the nonstationarity of the ongoing activity, a possible approach is to model the propagation in time of the probability density. In the case of Gaussian processes, a common solution is the Kalman filter model, which can be considered a discrete time special case of the Fokker-Planck equation (Astrom, 1970), and where both the mean and covariances functions are updated as time evolves, given new observations or measurements. In this case, the diffusion (propagation) is purely linear and the density function evolves as a Gaussian pulse that translates, spreads and is reinforced, remaining Gaussian throughout. The random component of the dynamical model leads to spreading – increasing uncertainty – while the deterministic component causes the density to drift. The effect of an external observation $z(t)$ is to superimpose a reactive effect on the diffusion, such that the density then tends to peak in the vicinity of the observation (Isard & Blake, 1998).

Here we chose a simpler, but perhaps more efficient and accurate approach when trying to capture the current statistical structure underlying a time series ensemble. The ongoing activity is treated to be locally WSS, i.e. over short time windows (e.g. 50 ms). In this way, instead of directly modeling the time evolution of the probability density, an entirely

new model is generated for every particular short time window centered at time t. Fast changes in brain dynamics and processes can then be tracked by the use of a moving time window. More specifically, for each time window centered at time t, an appropriate model order P is found together with a new set of coefficient matrices $\mathbf{A}_p$, for $p = 1,2,\ldots,P$. In other words, now we have a time dependent model expressed in terms of $\mathbf{A}_p(t)$, for $p = 1,2,\ldots,P(t)$, where the argument t refers to time where the moving time window is centered. This approach may be more reliable since by starting an entirely new model for every time window, the estimation is not affected by inaccuracies being propagated from the estimation at previous times. The moving window approach is referred to as Adaptive Multivariate Auto-Regressive (AMVAR) modeling. Extensive theoretical analysis, model validation and methods of assessment of statistical significance through bootstrap methods, on the same experimental data employed in this dissertation, has been presented in (Ding, Bressler, Yang, & Liang, 2000; Kaminski, Ding, Truccolo, & Bressler, 2001). Furthermore, for linear Gaussian processes and given an appropriate AMVAR model, the conditional joint probability

$$p(\eta(k),\eta(k-1),\ldots,\eta(k-p),\eta(k-p-1),\ldots|Event) \equiv p(\eta(k),\eta(k-1),\ldots,\eta(k-p),\eta(k-p-1),\ldots|t_{Event}) \quad (1.43)$$

can be completely specified from the knowledge of $\mathbf{A}_p(t_{Event})$ and $P(t_{Event})$. The above equation relates then to our original conditional probability in Equation (1.1).

The derived spectral quantities from the model are now time functions expressed as:

$$\Gamma_m(f,t) = S_{mm}(f,t), \quad (1.44)$$

$$C_{mn}(f,t)=\frac{|S_{mn}(f,t)|^2}{S_{mm}(f,t)S_{nn}(f,t)}, \tag{1.45}$$

$$D_{mn}(f,t)=|H_{mn}(f,t)|^2. \tag{1.46}$$

A comparison between the MVAR and a nonparametric high-resolution multitaper spectrogram is given in Figure 1-4. The AMVAR model and the nonparametric estimation were computed on the ensemble mean-subtracted time series from a parietal site. The spectral structure captured by the two methods agrees in general, while the AMVAR provides higher temporal resolution. The model order was five and the moving time window was 50 ms long. The multitaper estimation was computed on a single time window 600 ms long.

Two more aspects deserve mention when modeling cortical ongoing activity: continuous time models and mixture of deterministic and stochastic models. The fact that under appropriate conditions a continuous time linear dynamical model that preserves the same spectral structure can be derived from the fitted autoregressive model, gives yet another useful application of this type of modeling (Franaszczuk & Blinowska, 1985). For illustration, consider the univariate system

$$x(z) = H(z)\varepsilon(z). \tag{1.47}$$

The transfer function, derived from the AR modeling, is given by:

$$H(z)=\frac{1}{1-\sum_{p=1}^{P}a_p z^{-p}}=\frac{1}{-\sum_{p=0}^{P}a_p z^{-p}}, \tag{1.48}$$

with $a_0 = -1$. Multiplying both numerator and denominator by z^P:

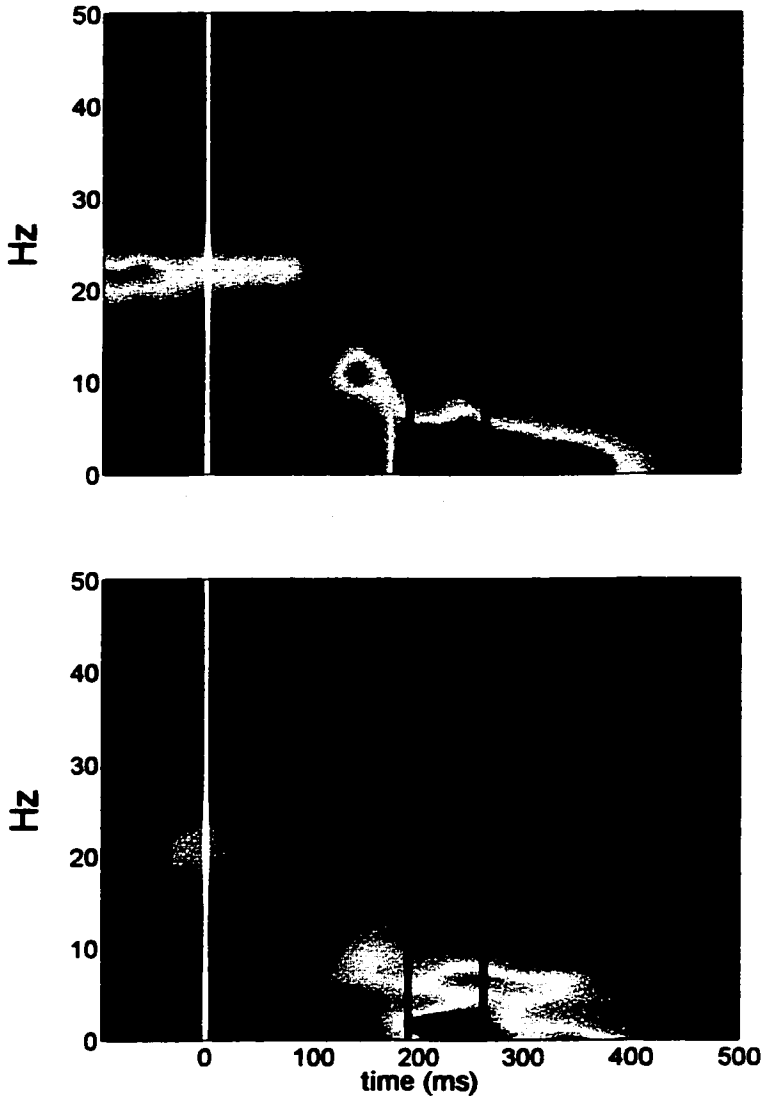

Figure 1-4: Comparison between AMVAR spectral estimation and non-parametric high-resolution multitaper spectrogram. The power spectrum density time function (top plot) for a parietal LFP activity was computed from a adaptive multivariate autoregressive model fitted on the whole LFP data ensemble (15 channels recorded intracortically). The model order was 5 (sampling interval 5 ms) and the moving window was 50 ms long (shifted by one data point at time). In the bottom plot, the Multitaper power spectrum computed on the single channel data from the same parietal site over the whole 600 ms time segment (no moving time window). The time-frequency estimation is obtained by projecting the Fourier components onto temporal eigenfunctions given by the Slepian functions themselves (Thomson, 2000). Notice that the AMVAR estimation captures basically the same spectral structure while allowing a better temporal resolution. Power is given in arbitrary units. The colored vertical lines follow the same convention as in the Figure 1-1. Ensemble size: 222 trials.

$$H(z)=\frac{z^{P}}{-\sum_{p=0}^{P}a_{p}z^{P-p}}. \tag{1.49}$$

Further, if $\beta_1,\beta_2,\ldots,\beta_P$ are the P roots of the polynomial $-\sum_{p=0}^{P}a_{p}z^{P-p}$, then after factorization of the denominator, $H(z)$ is:

$$H(z)=\frac{z^{P}}{-\prod_{p=1}^{P}(z-\beta_{p})}. \tag{1.50}$$

If we assume that there are only single poles of $H(z)$:

$$H(z)=\sum_{p=1}^{P}c_{p}\left(\frac{z}{z-\beta_{p}}\right), \tag{1.51}$$

where c_p are the single pole coefficients, obtained according to:

$$c_{p}=\lim_{z\to\beta_{p}}\frac{(z-\beta_{p})H(z)}{z}. \tag{1.52}$$

Employing the inverse Z-transform $\mathcal{Z}^{-1}$, the impulse response of the system (1.47) is expressed as:

$$h(t)=\mathcal{Z}^{-1}[H(z)]=\sum_{p=1}^{P}c_{p}e^{t\log\beta_{p}}. \tag{1.53}$$

According to Shannon's sampling theorem, if the sampling interval Δt has been appropriately chosen, the continuous time impulse response can be recovered from the discrete time one. The transfer function in continuous time can then be obtained through the Laplace transform:

$$H(s) = \mathcal{L}[\sum_{p=1}^{P} c_p e^{\frac{t}{\Delta t}\log\beta_p}] = \mathcal{L}[\sum_{p=1}^{P} c_p e^{\alpha_p t}]$$

$$H(s) = \sum_{p=1}^{P} c_p \frac{1}{s-\alpha_p}, \tag{1.54}$$

or in a rational form:

$$H(s) = \frac{1}{c_P s^P + \ldots + c_1 s + c_0}. \tag{1.55}$$

Given that the variable s corresponds to the operator $\frac{d}{dt}$, the model $x(s) = H(s)\varepsilon(s)$ can now be written as a P^{th} order differential equation:

$$c_P \frac{d^P x(t)}{dt^P} + \ldots + c_1 \frac{dx(t)}{dt} x(t) + c_0 x(t) = \varepsilon(t). \tag{1.56}$$

A model of spontaneous cortical ongoing activity obtained in this way usually corresponds to oscillatory damped modes with stochastic excitations (Franaszczuk & Blinowska, 1985). Similar models where obtained by Freeman were the direct experimental measure of the impulse responses of the neuronal populations were employed to derive purely deterministic differential equation models (Freeman, 1975). If the power spectrum of the time series contains many peaks, i.e. many characteristic oscillatory band activities, the transfer function can be decomposed into several transfer functions, each one accounting for the specific rhythms or time scales in the neural activity, perhaps capturing more closely the dynamics of very localized neural populations. As a result, the decomposed transfer functions may lead to 2nd order differential equations.

As a last topic, the problem of mixture of deterministic and stochastic models is briefly addressed. It could be thought that the apparent unpredictable nature of the spontaneous cortical activity should be considered more in terms of lower dimensional chaotic dynamics than in terms of a stochastic process. From a data analysis perspective, this leads to the difficulty of dealing with the modeling of time series generated by deterministic dynamics mixed with unpredictable components. This uncertainty can easily arise from measurement noise or from the contribution of unknown or not modeled factors. The difficulty resides in the fact that is practically impossible to decide on the specific nature of the deterministic component given only the recorded time series. For illustration, consider a time series $\mathbf{X}(t)$ obtained by appropriately sampling the following stochastic dynamical system:

$$\frac{d\mathbf{X}(t)}{dt} = F(\mathbf{X}(t)) + \xi(t), \tag{1.57}$$

where $F(\mathbf{X}(t))$ provides the dynamics of the Lorenz system, and $\xi(t)$ is an unpredictable component modeled as a white zero-mean Gaussian process, with diagonal covariance matrix $W_{ii} = \sigma_{\xi}^2$. For $\sigma_{\xi}^2 = 0$, a Markov process representation of the ensemble time series would require infinite order. That is because dynamical chaos has infinite memory, i.e., the mutual information between a state at time t and any future states given at $t+p$, i.e. temporal interdependencies, does not decrease but remains constant as p increases (Deco & Schurmann, 2000). However, as the variance of the white process incrementally increases, the same mutual information will decrease and tend to zero as p increases. If the contribution of the unpredictable component is large enough, a low order Markov

process will account for the statistical structure in the time series. At this point, there is no way of deciding, without specific additional information, if subjacent dynamical process generating the time series is a deterministic chaotic system or not. Once again, one of the central themes of this dissertation is played out here, namely, the role of signal-to-noise ratio on the estimation of interdependencies, in this case, temporal interdependencies.

Chapter 2

Sources of Event-Related Transients in Interdependence Measures

Chapter 1 has invoked two major sources of fast (~ 100 ms) event-related transients in interdependence measures computed on zero-mean or residual time series: nonlinear gain modulation and trial-to-trial variability of the event phase-locked responses. These two possibilities are now explored in detail with analytical and simulation tools.

2.1 Nonlinear Gain Modulation Effect

Dynamical models of neuronal population activities (Eeckman & Freeman, 1991b; Freeman, 1975) relate the field potentials and the pulse density (the number of

spikes per unit volume) in a local neuronal population by a nonlinear function of a sigmoid type. A resulting property of networks of such populations is the dynamic modulation of gain or effective connectivity by the network's mean activity level. This modulation can affect both the local population properties and the interactions among populations. First, transient changes in the population's mean level of activity, like that produced by stimulus-evoked responses, can place the system at regions of different gain levels of their output sigmoid function, resulting in amplification or attenuation of the ongoing activity measured by the transient changes in the time course of the ensemble variance and power spectral density time functions of the AERP-subtracted time series or residuals. Second, the residual time series from different populations, through the transient modulation of their shared variances, can exhibit an event-related modulation of their statistical interdependence measured by the fast transients in the time course of the cross-correlation (or coherence) time functions. Besides being a natural consequence in models of sigmoidally coupled neuronal populations, a possible functional relevance of this effect can be hypothesized in terms of slow spatio-temporal patterns of activity (say in the alpha range) determining the mean level of activity and subsequently soft-wiring the connectivity between neural components. The result would be a mechanism for dynamic functional connectivity in either broad or specific frequency bands (say gamma band).

This phenomenon is analyzed in the context of a model of excitatory and inhibitory neuronal populations coupled through sigmoid functions (Freeman, 1975). For

illustration, consider the following model of two interacting cortical columns containing excitatory and inhibitory neuronal populations:

$$\begin{aligned}
&\frac{d^2x_1}{dt^2}+(a+b)\frac{dx_1}{dt}+abx_1 = -k_3Q(y_1,Q_m)+k_1Q(x_2,Q_m)+\varepsilon_1(t)+I_1(t)\\
&\frac{d^2y_1}{dt^2}+(a+b)\frac{dy_1}{dt}+aby_1 = k_2Q(x_1,Q_m)+\varepsilon_2(t)\\
&\frac{d^2x_2}{dt^2}+(a+b)\frac{dx_2}{dt}+abx_2 = -k_3Q(y_2,Q_m)+k_1Q(x_1,Q_m)+\varepsilon_3(t)+I_2(t)\\
&\frac{d^2y_2}{dt^2}+(a+b)\frac{dy_2}{dt}+aby_2 = k_2Q(x_2,Q_m)+\varepsilon_4(t)
\end{aligned} \qquad (2.1)$$

Here $x(t)$ and $y(t)$ are the local field potentials of the excitatory and inhibitory populations, respectively, a and b are rate constants, and k gives the coupling gain between the populations. The spontaneous or ongoing activity is modeled by the introduction of the stochastic terms $\varepsilon(t)$, corresponding to white zero mean Gaussian process with variance σ_ε^2. $I(t)$ denotes the external input to the populations, employed to introduce fluctuations in the mean level of activity. The values of $Q(x,Q_m)$ and $Q(y,Q_m)$ represent pulse densities converted from x and y through a sigmoid function, with output gain or slope modulated by the parameter Q_m. Commonly, modulation in excitability or gain of the neural tissue resulting from attentional effects or other neuromodulatory factors is implemented in terms of variations in Q_m. Here we are considering a much faster and transient type of modulation in excitability that does not resort to this type of neuromodulatory effect, but relies exclusive on the nonlinear nature of the coupling function and on fluctuations in the mean level of activity. For this reason Q_m is assumed to be constant and it is henceforth dropped from the notation. An extensive stability and

bifurcation analysis of this system is provided in the Appendix for the sake of completeness.

For illustration, consider the symmetric condition $I_1(t) = I_2(t) \approx I$, with new fixed points after the input onset $x_1^* = x_2^*$ and $y_1^* = y_2^*$. The linearized system can then be rewritten:

$$\begin{aligned}
&\frac{d^2x_1}{dt^2} + (a+b)\frac{dx_1}{dt} + abx_1 = -K_3y_1 + K_1x_2 + \varepsilon_1(t) \\
&\frac{d^2y_1}{dt^2} + (a+b)\frac{dy_1}{dt} + aby_1 = K_2x_1 + \varepsilon_2(t) \\
&\frac{d^2x_2}{dt^2} + (a+b)\frac{dx_2}{dt} + abx_2 = -K_3y_2 + K_1x_1 + \varepsilon_3(t) \\
&\frac{d^2y_2}{dt^2} + (a+b)\frac{dy_2}{dt} + aby_2 = K_2x_2 + \varepsilon_4(t)
\end{aligned} \quad , \tag{2.2}$$

with new coupling strengths K (effective connectivity) defined as:

$$\begin{aligned}
K_1 &= k_1Q'(x_1^*) = k_1Q'(x_2^*) \\
K_2 &= k_2Q'(x_1^*) = k_2Q'(x_2^*). \\
K_3 &= k_3Q'(y_1^*) = k_3Q'(y_2^*)
\end{aligned} \tag{2.3}$$

Let $H(f) = \dfrac{1}{-\omega^2 + i\omega(a+b) + ab}$ and $\omega = 2\pi f$. After the Fourier Transform of Equation (2.2), we have:

$$\begin{aligned}
x_1(f) &= H(f)\left[K_1x_2(f) - K_3y_1(f) + \varepsilon_1(f)\right] \\
y_1(f) &= H(f)\left[K_2x_1(f) + \varepsilon_3(f)\right] \\
x_2(f) &= H(f)\left[K_1x_1(f) - K_3y_2(f) + \varepsilon_2(f)\right] \\
y_2(f) &= H(f)\left[K_2x_2(f) + \varepsilon_4(f)\right]
\end{aligned} \tag{2.4}$$

We are interested in obtaining the auto and cross-power density functions of x_1 and x_2 and examining how they are affected by variations in the sigmoid slope $Q'(v^*)$, around a

certain operational point ν^* dependent on the mean level activity. Substitution of $y_1(f)$ and $y_2(f)$ into the expressions for $x_1(f)$ and $x_2(f)$ results in:

$$\begin{aligned} x_1(f) &= H(f)\{K_1 x_2(f) - K_3 H(f)[K_2 x_1(f) + \varepsilon_2(f)] + \varepsilon_1(f)\} \\ x_2(f) &= H(f)\{K_1 x_1(f) - K_3 H(f)[K_2 x_2(f) + \varepsilon_4(f)] + \varepsilon_3(f)\} \end{aligned}, \tag{2.5}$$

or:

$$\begin{aligned} x_1(f) &= [1 + K_2 K_3 H^2(f)]^{-1} H(f)\{K_1 x_2(f) - K_3 H(f)\varepsilon_2(f) + \varepsilon_1(f)\} \\ x_2(f) &= [1 + K_2 K_3 H^2(f)]^{-1} H(f)\{K_1 x_1(f) - K_3 H(f)\varepsilon_4(f) + \varepsilon_3(f)\} \end{aligned}. \tag{2.6}$$

After substituting $x_2(f)$ into the expression for $x_1(f)$, and letting:

$$A(f) = \frac{H(f) + K_2 K_3 H^3(f)}{[1 + K_2 K_3 H^2(f)]^2 - K_1^2 H^2(f)}, \tag{2.7}$$

$$B(f) = \frac{K_3 H^2(f) + K_2 K_3^2 H^4(f)}{[1 + K_2 K_3 H^2(f)]^2 - K_1^2 H^2(f)}, \tag{2.8}$$

$$C(f) = \frac{K_1 H^2(f)}{[1 + K_2 K_3 H^2(f)]^2 - K_1^2 H^2(f)}, \tag{2.9}$$

$$D(f) = \frac{K_1 K_3 H^3(f)}{[1 + K_2 K_3 H^2(f)]^2 - K_1^2 H^2(f)}, \tag{2.10}$$

we obtain:

$$x_1(f) = A(f)\varepsilon_1(f) - B(f)\varepsilon_2(f) + C(f)\varepsilon_3(f) - D(f)\varepsilon_4(f), \tag{2.11}$$

and similarly,

$$x_2(f) = A(f)\varepsilon_3(f) - B(f)\varepsilon_4(f) + C(f)\varepsilon_1(f) - D(f)\varepsilon_2(f). \tag{2.12}$$

The power spectrum density of $x_1(t)$ and $x_2(t)$ are then respectively:

$$\Gamma_{x_1}(f) = < x_1(f) x_1^*(f) > = |A(f)|^2 \sigma_{\varepsilon_1}^2 + |B(f)|^2 \sigma_{\varepsilon_2}^2 + |C(f)|^2 \sigma_{\varepsilon_3}^2 + |D(f)|^2 \sigma_{\varepsilon_4}^2, \tag{2.13}$$

and

$$\Gamma_{x_2}(f) = < x_2(f)x_2^*(f) > = |A(f)|^2 \sigma_{\varepsilon_3}^2 + |B(f)|^2 \sigma_{\varepsilon_4}^2 + |C(f)|^2 \sigma_{\varepsilon_1}^2 + |D(f)|^2 \sigma_{\varepsilon_2}^2. \quad (2.14)$$

The cross-power spectrum density:

$$\begin{aligned} \Gamma_{x_1x_2}(f) &= < x_1(f)x_2^*(f) > \\ &= A(f)C^*(f)\sigma_{\varepsilon_1}^2 + C(f)A^*(f)\sigma_{\varepsilon_3}^2 + B(f)D^*(f)\sigma_{\varepsilon_2}^2 + D(f)B^*(f)\sigma_{\varepsilon_4}^2 \end{aligned}. \quad (2.15)$$

The coherence then becomes:

$$C_{x_1x_2}(f) = \frac{|\Gamma_{x_1x_2}(f)|}{\sqrt{\Gamma_{x_1}(f)\Gamma_{x_2}(f)}}. \quad (2.16)$$

To investigate the dependence of the derived power and coherence spectral density functions on the effective coupling strength K between the model columns, we fix the couplings K_2 and K_3 and vary K_1, the coupling between the two excitatory neuronal populations. Variation in K_1 represents variations in the sigmoid slope, since $K_1 = k_1 Q'(x_1^*) = k_1 Q'(x_2^*)$, which is dependent on the operational point set by the location of the new fixed points x_1^* and x_2^*, which themselves will depend on the level of the input I. Figure 2-1 show the power and coherence as a function of frequency and of K_1. Not surprisingly, the power and coherence clearly increase with increasing coupling. Overall the system behaves as combination of low pass filter, with a characteristic resonance in the gamma band, as seen by the peak around 40 Hz both in the power and coherence functions. This property results from the choice of values for the rate constants a and b, originally specified to account for the gamma oscillations in neuronal population in the olfactory system.

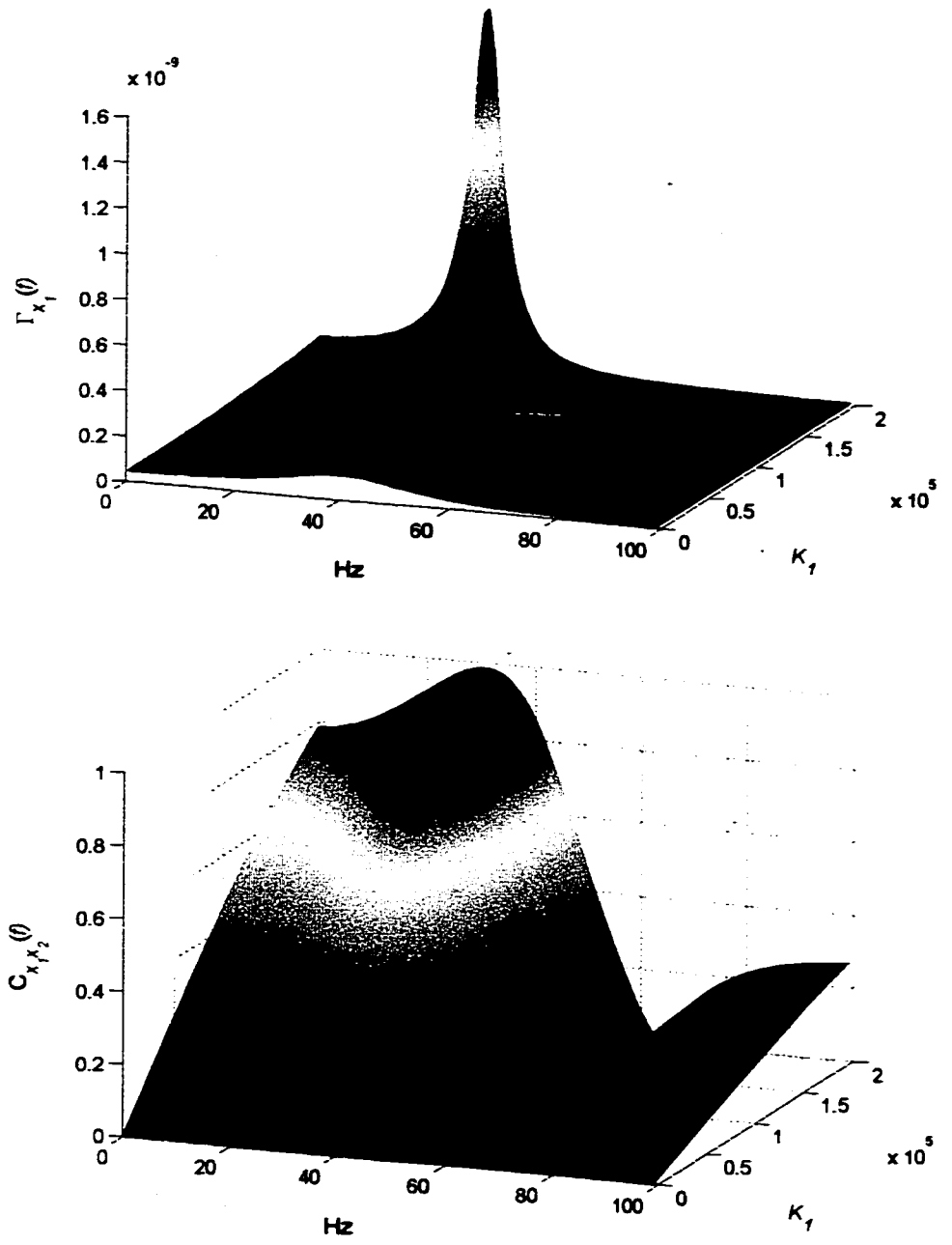

Figure 2-1: Power and coherence spectrum density dependence on effective coupling strength K_1. Power and coherence for the LFPs of neuronal populations x_1 and x_2 from the two columns linearized model are computed according to derived equations (see text). The coupling strength K_1 is dependent on the slope of the sigmoid output function, which itself depends on the level of LFP activity. As K_1 increases, both power and coherence increase. Model parameters: $K_2 = 10^5$, $K_3 = 4 \times 10^5$, $\sigma_\varepsilon = 1$, $a = 220/s$, $b = 720/s$.

Next we investigate the effect in the original nonlinear system and in the context the aimed situation of a transient fluctuation in the mean level of the activity of the neuronal populations, here mimicked by the effect of an external sinusoidal input. Transmission delays are also included, such that the system reads now:

$$\begin{aligned}
&\frac{d^2x_1}{dt^2}+(a+b)\frac{dx_1}{dt}+abx_1 = -k_3Q(y_1,Q_m)+k_1Q(x_2(t-\tau_{12}),Q_m)+\varepsilon_1(t)+I_1(t)\\
&\frac{d^2y_1}{dt^2}+(a+b)\frac{dy_1}{dt}+aby_1 = k_2Q(x_1,Q_m)+\varepsilon_2(t)\\
&\frac{d^2x_2}{dt^2}+(a+b)\frac{dx_2}{dt}+abx_2 = -k_3Q(y_2,Q_m)+k_1Q(x_1(t-\tau_{21}),Q_m)+\varepsilon_3(t)+I_2(t)\\
&\frac{d^2y_2}{dt^2}+(a+b)\frac{dy_2}{dt}+aby_2 = k_2Q(x_2,Q_m)+\varepsilon_4(t)
\end{aligned} \quad , \quad (2.17)$$

where

$$\begin{aligned}
&Q(V,Q_m) = Q_m(1-e^{-(e^V-1)/Q_m}) \text{ if } V > -u_o\\
&Q(V,Q_m) = -1 \text{ if } V \le -u_o \quad .\\
&u_o = -\ln(1+\ln(1+\frac{1}{Q_m}))
\end{aligned} \quad (2.18)$$

The above system is simulated for a network whose activity is concentrated in the gamma band (gamma network). The input to the two excitatory populations is given by a scaled single cycle sine, of period equal to 100 ms. The power and coherence spectrum (Figure 2-2) are computed on the AERP-subtracted time series, i.e., $\xi_m(t) = x_{mr}(t) - \langle x_{mr}(t)\rangle_r$. The power spectrum and coherence are respectively $\Gamma_{x_m}(f) \equiv \Gamma_{\xi_m}(f)$ and $C_{x_m x_n}(f) \equiv C_{\xi_m \xi_n}(f)$. A single AMVAR model is computed for the

time series ensemble of the variables $x_1(t)$ and $x_2(t)$. The ensemble mean activity (AERP) is shown in Figure 2-3. The same type of transient increase in power and coherence observed for the gamma-network can also be obtained in networks with oscillatory activity in other frequency bands, like alpha or beta, once the rate parameters a and b are adequately tuned (See Appendix for network frequency dependence on these parameters).

The main predictions of the nonlinear model are twofold. First, if in the ensemble of trials, the mean level of activity represented by the AERP tends to fluctuate beyond the near linear range of the sigmoid function, the ensemble variance time function $\sigma^2 = < [z_{mr}(t) - < z_{mr}(t) >_r]^2 >_r$ will show either peaks or valleys depending on if the LFP activity moves towards high or low slope regions of the sigmoid output function. These modulations in ensemble variance will obviously be accompanied by the same type of modulation in the power spectrum density time functions. A further prediction is that peaks or valleys in the variance should be accompanied by the same modulation in the cross-correlation time functions (see Equation (3.3). Similarly, the same relation should be observed for modulations in power spectrum and coherence time functions. It is difficult to predict how these modulations would relate to the waveform of the AERP, since it is not known which regions of the AERP correspond, on average, to a region of high slope of the sigmoid function. For example, if a positive extremum of the AERP relates to a region of high gain, peaks in the mentioned statistical quantities should tend to coincide in time with the AERP positive extremum. Also, if the latter is the case, negative extremum of the AERP should relate to low region gain (on average) of the sigmoid

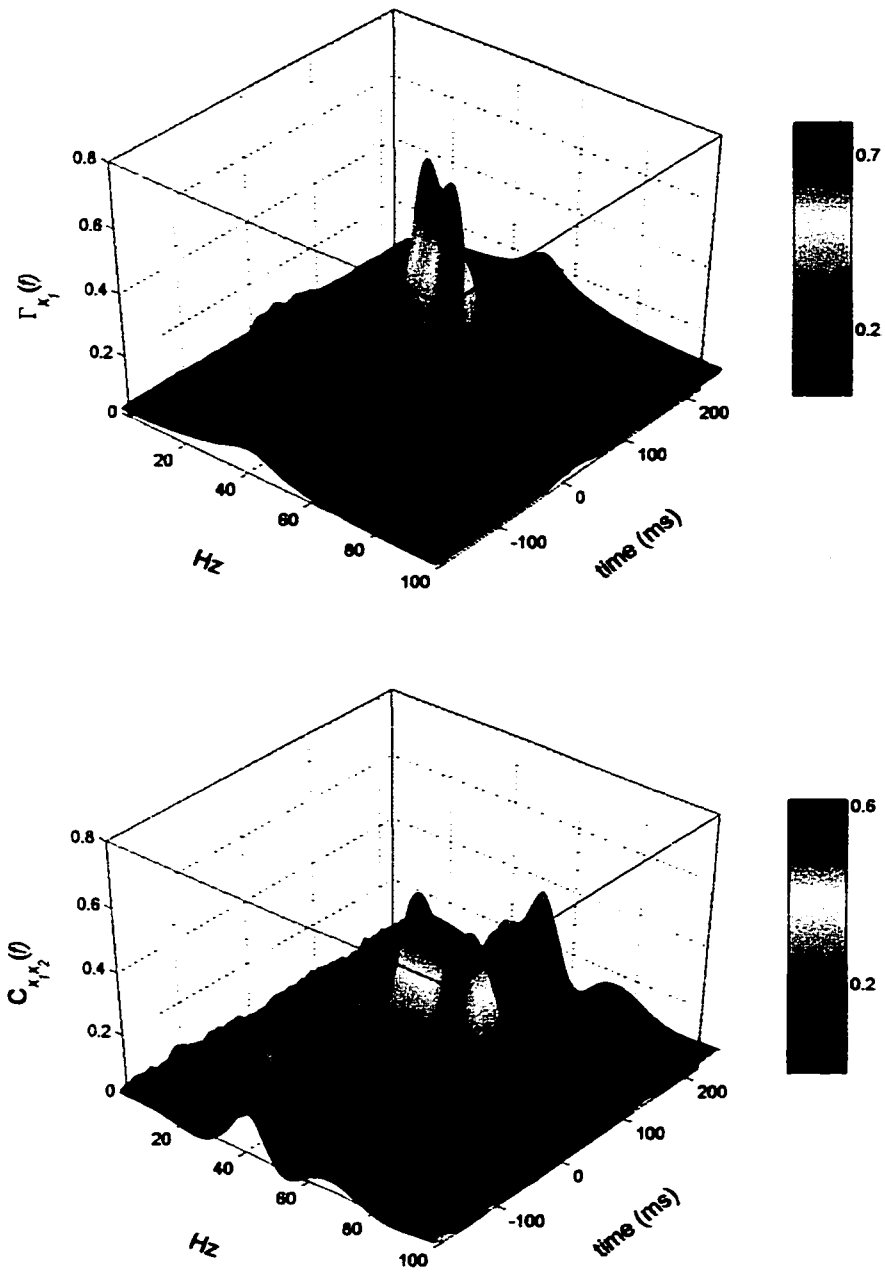

Figure 2-2: Transient nonlinear gain modulation of power and coherence time functions. The power (top plot) and coherence (bottom plot) are derived from a AMVAR model of the residual LFP time series of the neuronal populations x_1 and x_2, from the simulated two column system model (AMVAR order 5, sampling interval 5 ms, time window: 50 ms). The transient input consisted of a single cycle sine function (period of 100 ms), with amplitude 0.5. Stimulus onset is at time zero. Simultaneously there is an increase both in

power and coherence in the gamma band. Although the input generates a slow transient (~ 10 Hz) in the system response (see Figure 2-3), it does not contribute significantly to the power or coherence in the ~ 10 Hz range. This transient has been removed by the subtraction of the ensemble mean from each single trial. Model parameters: $k_1 = 0.05$, $K_2 = 0.1$, $K_3 = 0.4$, $\sigma_\varepsilon = 0.07$, $a = 220/s$, $b = 720/s$, $Q_m = 15$, $\tau_{12} = \tau_{21} = -10\,ms$. The deterministic part of the equations were integrated using the 4th order Runge-Kutta method, with linear interpolation for the delays. The stochastic part was integrated via the Euler method. Simulations were implemented in C. The power is given in arbitrary units. Ensemble size: 100 trials.

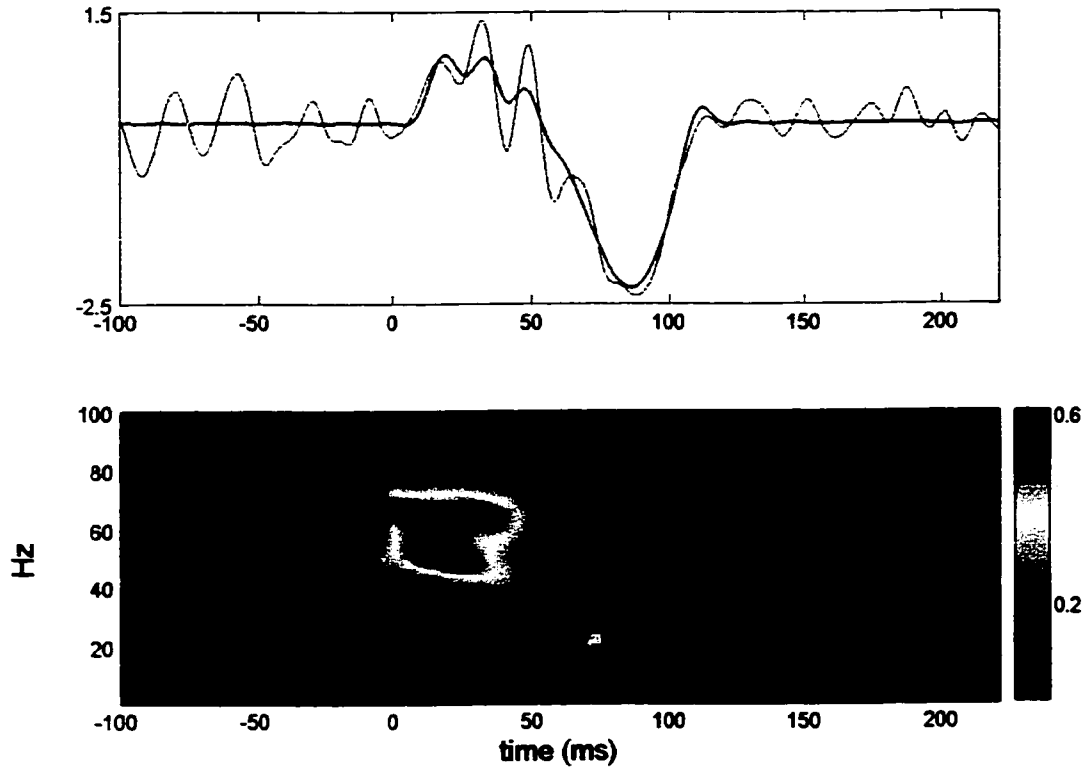

Figure 2-3: Single trial example, AERP and coherence for the same system and variables simulated in the previous Figure 2-2. The ensemble average or AERP (top plot, thick curve) captures the slow transient generated by the input and also some evoked activity on the gamma band that is phase-locked to the input onset. A single trial time series is shown by the thin curve. The spectral coherence increases in the positive phase of the input and decreases in the negative phase, returning to its background level after the transient is over.

function. Thus, valleys in the mentioned measures would tend to coincide with negative extrema of the AERP.

2.2 Trial-to-trial Nonstationarity of Event-Related Phase-Locked Responses

The signal-plus-noise (SPN) model can be formally expressed as follows:

$$z_r(t) = s(t) + \eta_r(t). \tag{2.19}$$

where $z_r(t)$ is the recorded cortical LFP at time t for the r^{th} trial, $s(t)$ is the stereotyped stimulus event-related response, and $\eta_r(t)$ is a zero-mean noise component, independent of the stimulus response, which represents ongoing noise activity. The stimulus onset is at $t = 0$. In practice, the average $< z_r(t) >_r$ taken over an ensemble of trials, $\{z_1(t), z_2(t), \ldots, z_R(t)\}$, is considered to be a consistent estimate of $s(t)$, i.e. $s(t) = < z_r(t) >_r$, and the ongoing noise is estimated as the residual remaining when the average is subtracted from the LFP time series, i.e.,

$$\xi_r(t) = z_r(t) - < z_r(t) >_r = \eta_r(t). \tag{2.20}$$

A more generic alternative model, the VSPN model, which takes into account the trial-to-trial variability in amplitude and latency of the event-related phase-locked activity, can be expressed by:

$$z_r(t) = a_r s(t - \tau_r) + \eta_r(t), \tag{2.21}$$

where the term a_r is a random variable corresponding to the amplitude of the event-related waveform and is assumed to be time independent for a given trial, and τ_r gives

the single trial latency of the response. The terms a_r, τ_r and $\eta_r(t)$ are assumed to be independent of each other.

2.2.1 The VSPN Model and its Predictions for the Single Channel Quantities

The effects of trial-to-trial variability in amplitude and latency in the VSPN model were analyzed separately. Their role in determining the relation between the ensemble mean and ensemble variance was central to understanding the post-stimulus behavior of power and interdependence measures.

2.2.1.1 Amplitude Variability

Considering the amplitude variability alone, the VSPN model simplifies to:

$$z_r(t) = a_r s(t) + \eta_r(t). \tag{2.22}$$

The ensemble average (AERP) based on this model becomes $< z_r(t) >_r = < a_r > s(t)$. When the AERP is subtracted from a single trial, the residual time series becomes:

$$\begin{aligned} \xi_r(t) &= a_r s(t) - < z_r(t) >_r + \eta_r(t) = (a_r - < a_r >) s(t) + \eta_r(t) \\ &= S_r(t) + \eta_r(t), \end{aligned} \tag{2.23}$$

which contains two components: the ongoing activity $\eta_r(t)$ and a component that is dependent on the AERP waveform:

$$S_r(t) = (a_r - < a_r >) s(t). \tag{2.24}$$

Thus, the VSPN model predicts that a stimulus phase-locked component will remain in the residual time series of LFP data after the AERP is subtracted from each trial.

The variance of $z_r(t)$ over the ensemble of trials at a given time t is equal to the ensemble variance of the residual $\xi_r(t)$, which from Equation (2.22) becomes:

$$\sigma^2(t) = <[\xi_r(t)]^2> = <[S_r(t)]^2> + <[\eta_r(t)]^2>$$

$$= \sigma_S^2(t) + \sigma_\eta^2(t), \qquad (2.25)$$

where the variance of the stimulus phase-locked component is given by:

$$\sigma_S^2(t) = <[a_r - <a_r>]^2> [s(t)]^2 \ . \qquad (2.26)$$

Since $\eta_r(t)$ is assumed to be stationary over the trial length, it follows that the variance of the residual has a time course resembling that of $[s(t)]^2$. In particular, the peaks of the variance function occur at the same times as the extrema of the AERP. Moreover, the power spectral density time function of the residual, $<|\xi_r(f,t)|^2>$, computed in a sliding time window centered at time t, will also be modulated according to the AERP shape at frequencies characteristic of the AERP waveform. Specifically, since $s(t)$ is often oscillatory with a very distinct main frequency, the quantity $<|\xi_r(f,t)|^2>$ at this frequency will be similarly modulated and should also exhibit a significant increase during the evoked response time period.

Since the amplitude of the phase-locked component, a_r, can be either greater or less than the average on any single trial, it follows that $(a_r - <a_r>)$ can take both positive and negative values. This effect is illustrated using simulated data in Figure 2-4. Therefore, the Fourier component $\xi(f)$ at the characteristic frequency f, exhibits a phase distribution having two modes that represent the positive and negative amplitudes, and thereby differ by the value of 180 degrees. By contrast, if the phase-locked component in

the residual were zero, as predicted by the SPN model, a uniform phase distribution would result.

2.2.1.2 Latency Variability

When only the latency variability is considered, the VSPN model simplifies to:

$$z_r(t) = s(t-\tau_r) + \eta_r(t). \tag{2.27}$$

To evaluate the effect of the right hand side of Equation (2.27), we use the Taylor expansion of $s(t-\tau_r)$ for small τ_r, which up to the first order is:

$$s(t-\tau_r) = s(t) - s'(t)\tau_r, \tag{2.28}$$

where $s'(t)$ is the first derivative with respect to t. The approximation to (2.27) then becomes:

$$z_r(t) = s(t) - s'(t)\tau_r + \eta_r(t). \tag{2.29}$$

By letting $<\tau_r> = 0$, the ensemble mean is then given by:

$$<z_r(t)>_r = s(t), \tag{2.30}$$

and the residual time series becomes:

$$\begin{aligned}\xi_r(t) &= s(t) - <z_r(t)>_r - s'(t)\tau_r + \eta_r(t) = -s'(t)\tau_r + \eta_r(t) \\ &= S_r(t) + \eta_r(t).\end{aligned} \tag{2.31}$$

Again the residual time series contain a component, $S_r(t) = -s'(t)\tau_r$, that is phase-locked to stimulus onset. From Equation (2.31) we obtain the ensemble variance as

$$\begin{aligned}\sigma^2(t) &= <[\xi_r(t)]^2>_r = <[\tau_r]^2>[s'(t)]^2 + <[\eta_r(t)]^2>_r \\ &= \sigma_S^2(t) + \sigma_\eta^2(t),\end{aligned} \tag{2.32}$$

where the variance of the phase-locked component is approximately:

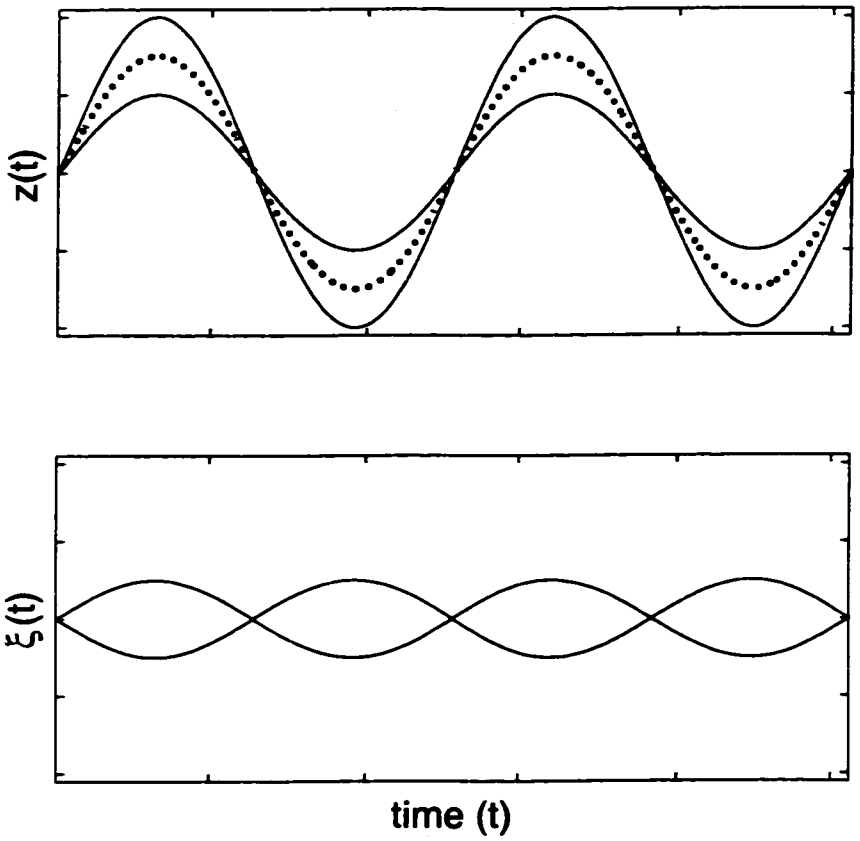

Figure 2-4: Simulation using a sine wave signal of variable amplitude (with independent noise set to zero). Top: two realizations of the signal having different amplitudes are shown by the solid curves, and the ensemble mean is shown by the dotted curve. Bottom: the residuals that result from subtracting the ensemble mean from each realization, the phase of the residuals being equal to either 0 or 180 degrees. This simulation illustrates that phase-locked components remain in the residual time series after subtraction of the ensemble mean.

$$\sigma_S^2(t) = <[\tau_r]^2 > [s'(t)]^2. \quad (2.33)$$

This equation shows that the ensemble variance is again modulated by the evoked response since it follows the square of the first derivative of the evoked response. Equation (2.33) allows us to predict the form of the variance function when the AERP is a damped sinusoid. That is because the derivative of a sinusoid has the property of being shifted by 90 degrees. Specifically, maximum values in variance should coincide in time with inflexion points of the AERP, and minimum values in variance with extrema of the AERP. As an example of this effect, we consider a simulated time series $z_r(t) = s(t - \tau_r)$, without added noise, and with the waveform given by $s(t - \tau_r) = H(t - \tau_r)\sin(\omega t - \phi_r)$, where $H(t)$ is a Hanning window, used to produce a waveform that mimics commonly observed AERPs, and $\phi_r = \tau_r \omega$. The signal $s(t)$ contains 2 full cycles of a sinusoidal oscillation with period T = 50 (arbitrary time unit), and is zero elsewhere. Figure 2-5 shows two examples where the single trial parameter ϕ_r is taken from uniform distributions over the following intervals: (1) [$-\pi/2$, $\pi/2$], and (2) [$-3\pi/4$, $3\pi/4$]. The results demonstrate that the main characteristics of the modulation of ensemble variance are well captured by the linear approximation in Equation (2.32) even for large latency variability.

2.2.1.3 Amplitude and Latency Variability

When both amplitude and latency vary from trial to trial, their interplay can have complicated effects. We use a simple example to illustrate the effect of increasing latency variability while maintaining a constant level of amplitude variability (Figure 2-6). For

small latency variability, the amplitude variability dominates, resulting in variance peaks that coincide with extrema of the AERP (top plot). As the latency variability increases, a point is reached where amplitude and latency effects are balanced, leading to a single peak in the variance function (middle plot). As latency increases even further (bottom plot), the characteristic effect of latency variability as described by the linear approximation in Equation (2.32). At the highest levels of latency variability (not shown), a single peaked or flat variance is again observed.

2.2.2 Implications of the VSPN Model for Statistical Interdependence Measures Involving Two Channels

Consider two channel recordings given by $z_{1r}(t) = a_{1r}s_1(t) + \eta_{1r}(t)$ and $z_{2r}(t) = a_{2r}s_2(t) + \eta_{2r}(t)$. The cross-correlation function between the two channels' time series, at time lag τ and time t is then given by:

$$C_{z_1 z_2}(\tau, t) = \frac{< S_{1r}(t)S_{2r}(t-\tau) >_r + < \eta_{1r}(t)\eta_{2r}(t-\tau) >_r}{\sqrt{< [S_{1r}(t)]^2 >_r + < [\eta_{1r}(t)]^2 >_r}\sqrt{< [S_{2r}(t-\tau)]^2 >_r + < [\eta_{2r}(t-\tau)]^2 >_r}}, \tag{2.34}$$

where $S_{1r}(t) = (a_{1r} - < a_{1r} >)s_1(t)$ and $S_{2r}(t) = (a_{2r} - < a_{2r} >)s_2(t)$. Without lost of generality we assume that the ongoing noise processes are approximately stationary, meaning that $< \eta_{1r}(t)\eta_{1r}(t-\tau) >_r$, $< [\eta_{1r}(t)]^2 >_r$ and $< [\eta_{1r}(t-\tau)]^2 >_r$ have no time dependence. That is, the cross-correlation function between the noise processes, $\eta_{1r}(t)$ and $\eta_{1r}(t-\tau)$, is only a function of the time lag τ and is time independent:

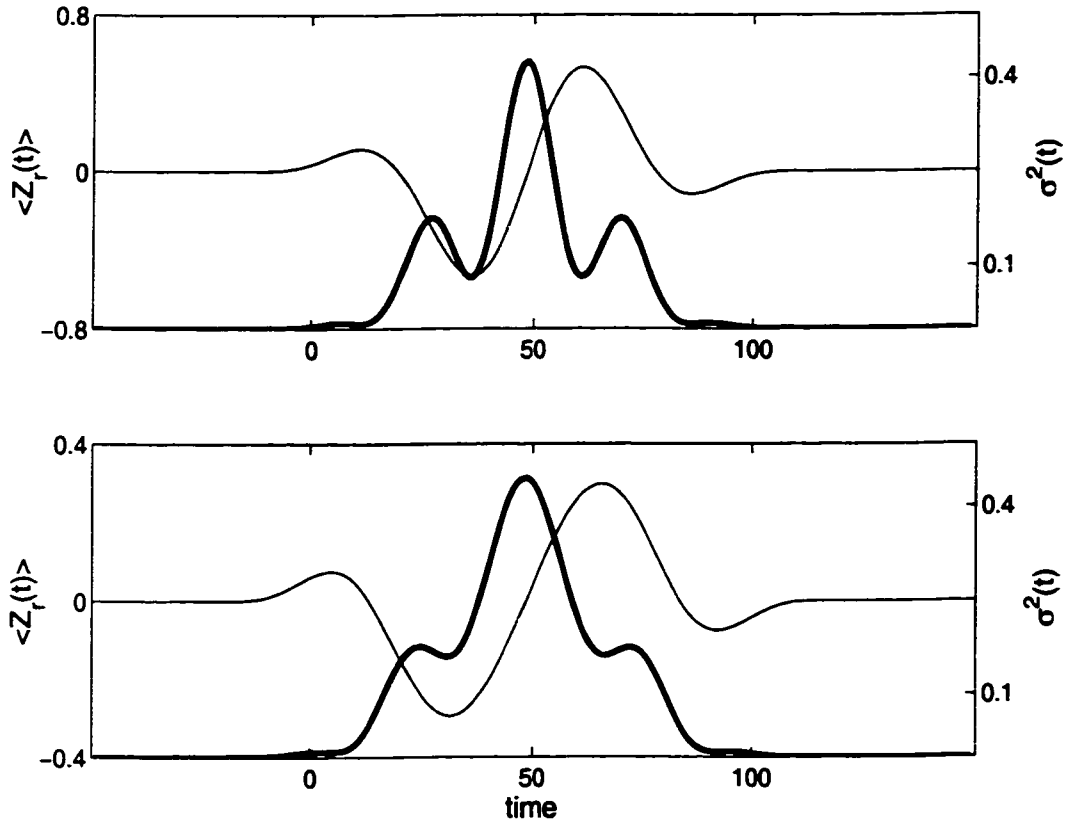

Figure 2-5: Relation between ensemble average, $< z_r(t) >_r$, and ensemble variance time function, $\sigma^2(t)$, for the case of trial-to-trial latency variability in simulated time series. The ensemble variance is shown by the thicker curve. The latency parameter expressed in terms of phase (see text) is uniformly distributed in the intervals $[-\pi/2, \pi/2]$ (top plot) and $[-3\pi/4, 3\pi/4]$ (bottom plot). The derived relations work well for this wide range of latency variability: peaks in variance coincide with inflexion points of the evoked response waveform, represented here by the ensemble average, and valleys in the variance coincide with extrema of the ensemble average. The simulated evoked signal $s(t)$ consisted of two full cycles of a sinusoidal waveform. Here the amplitude, variances and time are given in arbitrary units and the number of trials used was 2000.

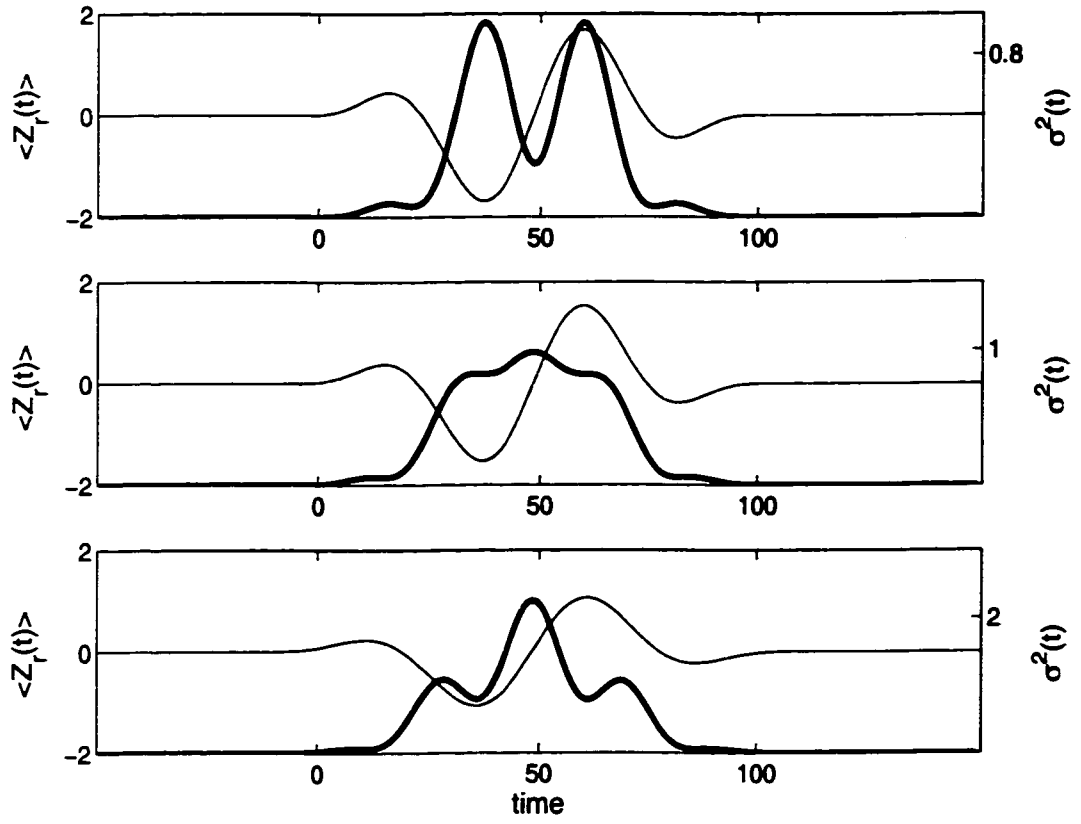

Figure 2-6: Relation between ensemble average, $< z_r(t) >_r$, and ensemble variance time function, $\sigma^2(t)$, for the case of trial-to-trial amplitude and latency variability in simulated time series. The same model as in Figure 2-5 was simulated and the ensemble variance is again shown by the thicker curve. Amplitudes were taken from a uniform distribution over the interval [0, 4]. The latency parameter expressed in terms of phase (see text) was uniformly distributed over the following intervals: $[-\pi/10, \pi/10]$ (top plot), $[-\pi/4, \pi/4]$ (middle plot) and $[-\pi/2, \pi/2]$ (bottom plot). When the amplitude variability dominates (top plot), the ensemble variance peaks coincide with extrema of the AERP. As the latency variability increases, amplitude and latency are balanced, resulting in a single peak being observed (middle plot). As latency variability increases even further, the variance peaks coincide with inflexion points of the AERP waveform, and variance valleys coincide with extrema of the AERP (bottom plot). A total of 2000 trials were employed. Amplitude, variances and time are given in arbitrary units.

$$C_{\eta_1\eta_2}(\tau,t) = C_{\eta_1\eta_2}(\tau) = <\eta_{1r}(t)\eta_{2r}(t-\tau)>_r / \sqrt{<[\eta_{1r}(t)]^2>_r <[\eta_{2r}(t-\tau)]^2>_r}\ .$$

The cross-correlation function between the two evoked signals, considering the case of trial-to-trial amplitude variability, is given by:

$$C_{S_1S_2}(\tau,t) = \frac{<(a_{1r}-<a_{1r}>)(a_{2r}-<a_{2r}>)> s_1(t)s_2(t-\tau)}{\sqrt{<[a_{1r}-<a_{1r}>]^2>}\sqrt{<[a_{2r}-<a_{2r}>]^2>}|s_1(t)||s_2(t-\tau)|} = \pm K, \quad (2.35)$$

for $s_1(t) \neq 0$, $s_2(t-\tau) \neq 0$ and where K is the correlation coefficient of the two random amplitudes. Although $|C_{S_1S_2}(\tau,t)|$ is a constant, $C_{S_1S_2}(\tau,t)$ may flip between positive and negative values of K, depending on the sign of the product $s_1(t)s_2(t-\tau)$ which may change in time for a given time lag. Thus, the very presence of the variable signals in the VSPN model already makes possible temporal modulations of the cross-correlation function between the two time series, $C_{z_1z_2}(\tau,t)$, albeit in a simple and very stereotypical way. The interplay between signals and noises can lead to more complex and significant temporal modulations of $C_{z_1z_2}(\tau,t)$, as we consider below.

Let *SNR* represent the signal-to-noise ratio, defined as $SNR_1(t) = \sigma^2_{S_1}(t)/\sigma^2_{\eta_1}(t)$ and $SNR_2(t-\tau) = \sigma^2_{S_2}(t-\tau)/\sigma^2_{\eta_2}(t-\tau)$. Dividing both the numerator and denominator by $\sigma_{\eta_1}(t)\sigma_{\eta_2}(t-\tau)$, and after algebraic manipulation, Equation (2.34) can be rewritten as:

$$C_{z_1z_2}(\tau,t) = \frac{\sqrt{SNR_1(t)SNR_2(t-\tau)}C_{S_1S_2}(\tau,t) + C_{\eta_1\eta_2}(\tau)}{\sqrt{SNR_1(t)+1}\sqrt{SNR_2(t-\tau)+1}}. \quad (2.36)$$

It is clear that complex temporal modulation can result depending on the time course of the signal-to-noise ratios. In particular, as the product of the signal-to-noise ratios,

$SNR_1(t)SNR_2(t-\tau)$, increases, the cross-correlation function approaches the evoked signals' cross-correlation function, i.e., $C_{z_1z_2}(\tau,t) \rightarrow C_{S_1S_2}(\tau,t)$. As seen in the previous section, fluctuations in the signal-to-noise ratios are expected to occur since the signals' variances have their time course modulated by the square of the AERP waveform. Two main characteristic cases showing this effect, referred to as the *time-dependent signal-to-noise ratio effect*, are discussed below.

Case 1. Transient increase[2] *in cross-correlation.* To illustrate this effect, consider the situation where for a given time lag τ we have that $C_{S_1S_2}(\tau,t) = K > C_{\eta_1\eta_2}(\tau,t)$ and that $SNR_1(t) \approx SNR_2(t-\tau)$. Thus the stimulus phase-locked components in the residual time series co-vary at a level greater than that of the ongoing noise. Equation (2.36) then becomes:

$$C_{z_1z_2}(\tau,t) = \frac{SNR_1(t)K + C_{\eta_1\eta_2}(\tau)}{SNR_1(t)+1}. \qquad (2.37)$$

This function is a monotonically increasing function of $SNR_1(t)$, and if we assume that $\sigma^2_{\eta_1}(t)$ is constant, $C_{z_1z_2}(\tau,t)$ will follow the profile of $\sigma^2_{S_1}(t)$. Therefore, peaks of the cross-correlation function will coincide with peaks in the variance of the stimulus phase-locked component, and hence, as demonstrated above, with extrema of the AERP. This predicted signature will be demonstrated in the analysis of LFP data that follows, suggesting that this case represents a realistic situation. The possibility will be discussed that cross-correlation increases that are actually due to signal-to-noise ratio variation

might be misconstrued as reflecting modulation in the level of synchronization between two neuronal populations.

An example of this effect is now provided, with two simulated time series described as: $z_{1r}(t) = a_r s(t) + \eta_{1r}(t)$ and $z_{2r}(t) = a_r s(t) + \eta_{2r}(t)$. The time series share a common signal that is given by $s(t) = H(t)\sin(\omega t)$, consisting of two cycles (period T = 50), existing from t = 0 to t = 100, and zero elsewhere (Figure 2-7). $H(t)$ is a Hanning window. The ongoing activities are modeled as zero-mean, unit-variance, Gaussian, white-noise processes. The single trial amplitudes a_r are the same for the two channels and are drawn from a uniform distribution over the interval [0, 4]. The zero-time-lag, point-by-point cross-correlation is computed according to (3.3). As expected, it peaks at the same times as the variance of the residuals (i.e., at the signal-to-noise ratio peaks) and oscillates at half the period of $s(t)$.

For situations where the main characteristic frequencies of $s_1(t)$ and $s_2(t)$ are significantly different, peaks in the cross-correlation function may no longer coincide precisely with peaks in variance, although as before, the same time-dependent signal-to-noise ratio effect will be observed, i.e. $C_{z_1 z_2}(\tau, t) \rightarrow C_{S_1 S_2}(\tau, t)$ as the signal-to-noise ratio increases (Figure 2-8).

Case 2. Transient decrease in cross-correlation. We next consider the situation where the stimulus phase-locked components $S_{1r}(t)$ and $S_{2r}(t-\tau)$ in the residual time

---

[2] We actually refer to "increase" or "decrease" of the absolute value of the cross-correlation function. Also, we refer to peaks in cross-correlation function in the sense of peaks in its absolute value.

series do not co-vary, i.e., $<(a_{1r}-<a_{1r}>)(a_{2r}-<a_{2r}>)>=0$, but the ongoing activities are correlated, i.e., $<\eta_{1r}(t)\eta_{1r}(t-\tau)>_r = C_{\eta_1\eta_2}(\tau)>0$. Assuming as before that the time dependence of $\sigma^2_{\eta_1}(t)$ and $\sigma^2_{\eta_2}(t-\tau)$ are weak, then the value of the observed cross-correlation $C_{z_1z_2}(\tau,t)$ will still be modulated by the time dependence of the product of the signal-to-noise ratios. As the signal-to-noise ratio product increases, the observed cross-correlation time function approaches that of the stimulus phase-locked components, which is zero. In practice, the resulting modulation could be misconstrued as a short transient desynchronization episode between two recorded neuronal populations (see Figure 2-9).

For frequencies *f* related to the main oscillatory components of the evoked signals, the same effects described in cases 1 and 2 will also be observed in the time-varying spectral coherence function:

$$C_{z_1z_2}(f,t)=\frac{\left|<S_{1r}(f,t)S^*_{2r}(f,t)>_r+<\eta_{1r}(f,t)\eta^*_{2r}(f,t)>_r\right|}{\sqrt{<|S_{1r}(f,t)|^2>_r+<|\eta_{1r}(f,t)|^2>_r}\sqrt{<|S_{2r}(f,t)|^2>_r+<|\eta_{2r}(f,t)|^2>_r}}, \tag{2.38}$$

which is computed in a sliding time window centered at time *t*. This can be easily verified considering the similarity between equations (2.38) and (2.34).

Even though we have only considered fluctuations in signal-to-noise ratio that originate from trial-to-trial amplitude variability of the evoked signals, latency variability will result in similar fluctuations, and consequently in temporal modulation of the same interdependence measures. For latency variability that is not too large, peaks in the signal-

to-noise ratios are expected to coincide with inflexion points of the AERP waveform, rather than with extrema as in the case of amplitude variability just described. For the more realistic case where both amplitude and latency variability occur, the combined effect will depend on which one dominates the modulation of ensemble variance, as exemplified in Figure 2-6.

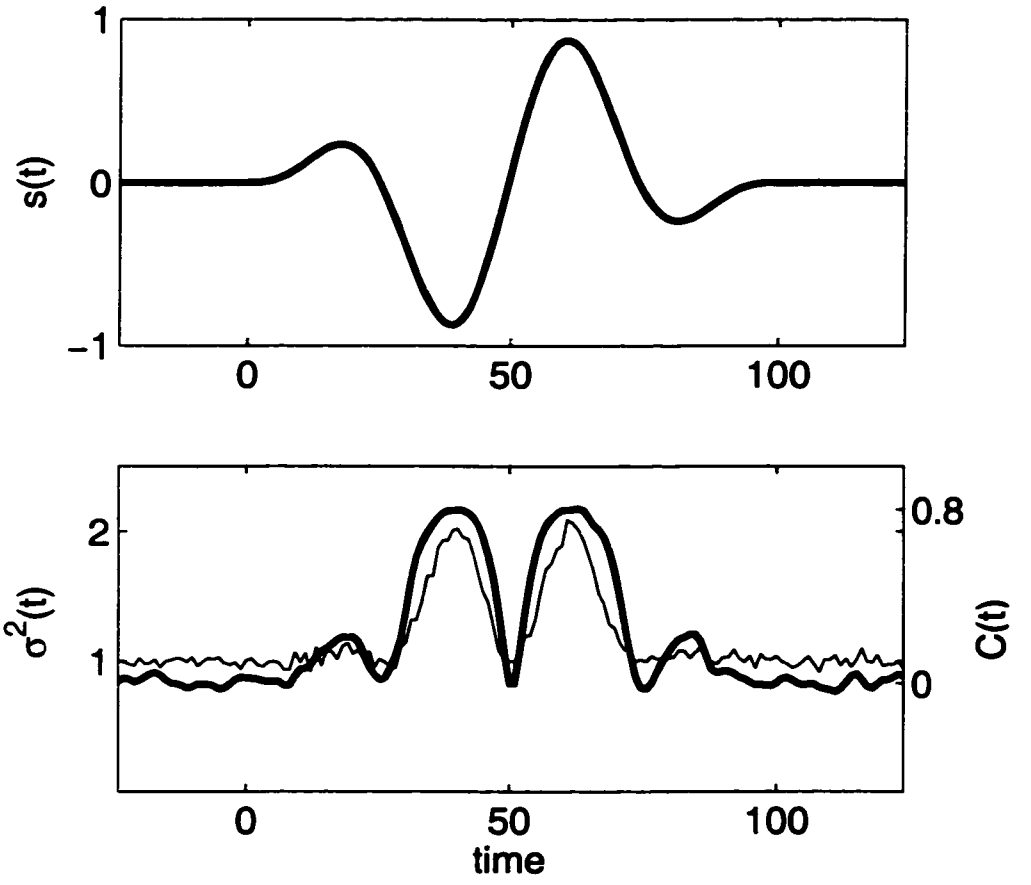

Figure 2-7: Time-dependent signal-to-noise ratio effect on the modulation of the cross-correlation time function: transient increase I. The two channels share a common evoked response *s(t)* (top plot) whose amplitude is drawn from a uniform distribution over the interval [0, 4]. Each channel's ongoing activity was modeled as independent zero mean and unit variance white Gaussian noise. In the bottom plot, the zero lag cross-correlation, *C(t)*, (thick curve) and the ensemble variance $\sigma^2(t)$ time functions are shown. The variance (and the signal-to-noise ratio) of the shared evoked component oscillates at ½

the period of the original sinusoidal component. As a consequence, the cross-correlation time function also oscillates with the same period, with peaks coinciding in time with peaks in the variance, or equivalently, with extrema of the evoked response. The ensemble consisted of 2000 simulated trials. Amplitude, variances and time are given in arbitrary units. The ensemble variance time functions of each channel are approximately the same and only one of them is shown.

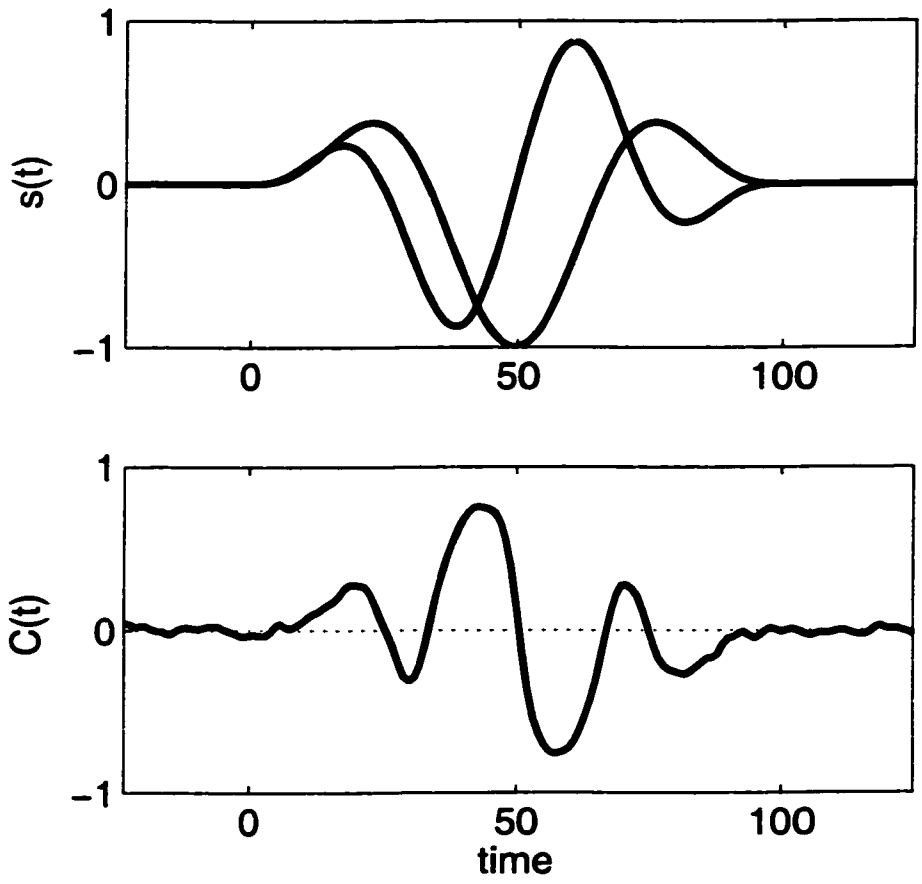

Figure 2-8: Time-dependent signal-to-noise ratio effect: transient increase II. Top: simulated evoked responses, *s(t)*, from two channels. Parameters for the evoked response waveforms, amplitudes and ongoing noise are the same as in Figure 2-7, except that the periods of the two evoked signals are slightly different. Bottom: zero lag cross-correlation *C(t)* (thick curve) computed from an ensemble of 2000 simulated trials composed of the evoked responses linearly combined with noise. The extrema in the cross-correlation function in general may not agree with extrema of *s(t)* or equivalently with peaks of the ensemble variance from each channel. Amplitudes and time are given in arbitrary units.

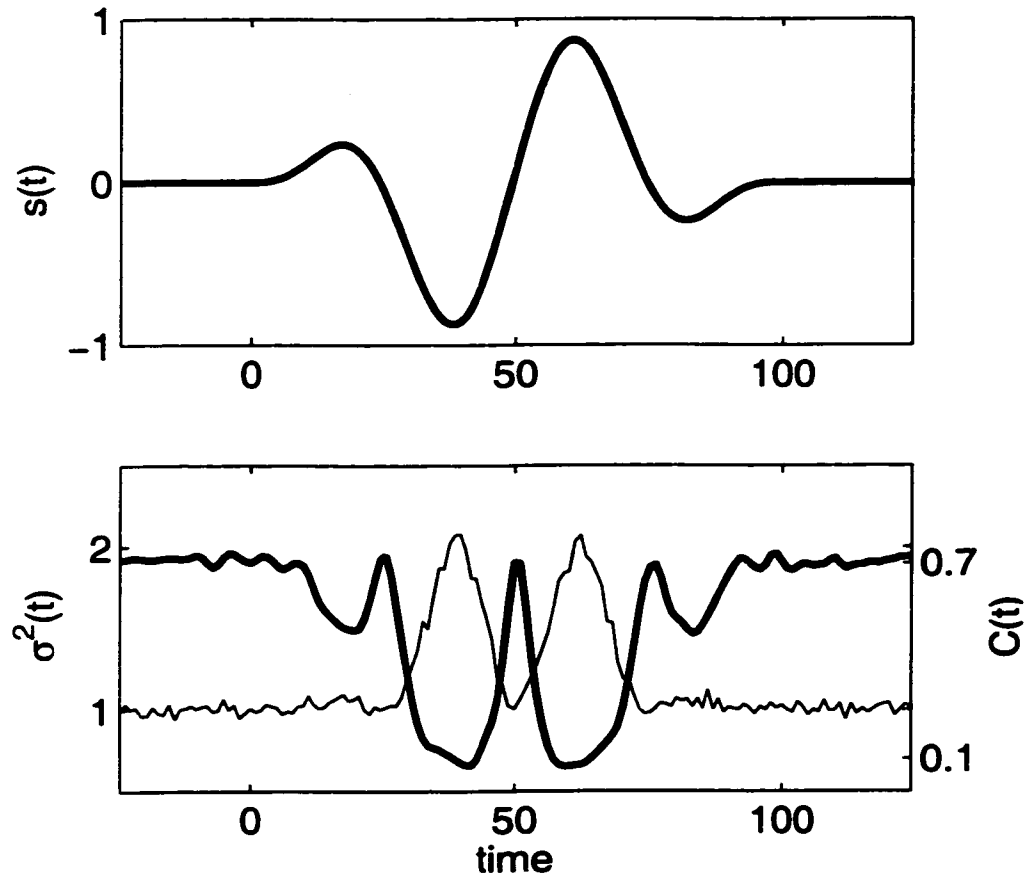

Figure 2-9: Time-dependent signal-to-noise ratio effect: transient decrease. Top: simulated evoked responses, $s(t)$, from two channels. Parameters for the evoked response waveform and amplitudes are the same as in Figure 2-7, except that the trial-to-trail amplitudes for the common evoked responses for each channel are independent, while the ongoing noise is now correlated. Bottom: ensemble variance $\sigma^2(t)$ and zero lag cross-correlation $C(t)$ time functions, computed from an ensemble of 2000 simulated trials composed of the evoked responses linearly combined with noise. The cross-correlation is shown by the thick curve. Here the signal-to-noise ratio effect appears as a transient decrease in cross-correlation that could resemble a transient desynchronization event. Peaks in the variance time function of the time series from each channel coincide with minima of the cross-correlation time function. The ensemble variance time functions of each channel are approximately the same and only one of them is shown.

Chapter 3

Event-Related Transients in Interdependence Measures: Experimental Results

The predicted effect on interdependence measures derived in Chapter 2 for both the nonlinear gain effect and for the trial-to-trial variability of cortical LFP responses are now investigated in an experimental data set involving a visuo-motor pattern discrimination task executed by monkeys.

3.1 Methods

3.1.1 Visuo-Motor Pattern Discrimination Task

Experiments were performed on four macaque monkeys (LU, GE, PE and TI) in the Laboratory of Neuropsychology at the National Institute of Mental Health. Animal

care was in accordance with institutional guidelines at the time. A detailed description of the experiment has previously been presented (Bressler, Coppola, & Nakamura, 1993). The monkeys performed a visual pattern discrimination task in which they sat facing a visual display screen. Each trial began with self-initiated depression of a hand lever with the preferred hand. Following a random interval from 0.5 to 1.2 s, a visual stimulus appeared on the screen for 100 ms. The stimulus was one of two pattern types: four dots arranged as a diamond or as a line. The monkey indicated its discrimination of line and diamond patterns by releasing (GO) or maintaining (NO-GO) pressure on the lever. The monkey received a water reward for correct GO performance. GO and NO-GO trials were randomly presented with equal probability in 1000-trial sessions. The contingency between stimulus pattern and response type was reversed across sessions. The present study focuses on the initial stages of the visual evoked response to presentation of the line pattern type. For each ensemble, the following numbers of trials were used: 2344, 888, 914 and 2053 for the monkeys LU, GE, PE and TI, respectively. For each monkey, the ensemble of trials was balanced for response type (GO and NO-GO).

3.1.2 Recording and Data Preprocessing

LFPs were sampled at 200 Hz from chronically implanted surface-to-depth transcortical bipolar electrodes at 11 to 15 cortical sites in the cerebral hemisphere contralateral to the preferred hand. Approximate localization of the electrodes in each monkey is shown in Figure 3-1. Each trial was recorded from approximately 115 ms pre-stimulus onset to 500 ms post-stimulus. The LFPs in each trial were linearly detrended,

60 Hz removed, and normalized to zero mean and unit variance (time normalization). After these preprocessing steps, the AERPs (ensemble averages) and ensemble variance time functions were computed, and constitute then normalized quantities. In addition, each ensemble of trials was normalized at each time point to zero ensemble mean and unit ensemble variance (ensemble normalization). The resulting residual time series were employed in the computation of the phase histograms, cross-correlations and spectral quantities described below.

3.1.3 AMVAR Spectral Estimation

Power and coherence time functions were estimated by the application of Adaptive Multivariate AutoRegressive (AMVAR) models to the residual time series (see Chapter 1, Section 1.2.2). Each single trial residual (600 ms long) was divided in 110 consecutive and overlapping (shifted by one data point) windows of 10 points (50 ms) each. A MVAR model of order 5 was adaptively computed for each successive window using the LWR algorithm (Haykin & Kesler, 1983). A spectral matrix was derived for each time window from the model coefficients, and then used as the basis for computing power and coherence spectra in the range of 1 to 100 Hz. The squared coherence values were used. The moving time windows then allowed the construction of power and coherence time functions for any frequency bin. AMVAR spectral estimation was employed instead of the traditional nonparametric approach of directly applying the Direct Fourier Transform (DFT) to the time series data because the DFT is an extremely biased estimate when used in such short time windows. The choice of 50 ms as the time

window duration came from previous experience with the application of AMVAR analysis to this data set (Ding, Bressler, Yang, & Liang, 2000). Overall, the use of short time windows in this paper was motivated by the need for tracking nonstationary transient cortical processes on a sub-second time scale. Particular attention was given to power and coherence at the 12 Hz spectral component. This had two main motivations. Following the VSPN hypothesis, the stimulus phase-locked component in the residual should have the same frequency content as that of the AERP. For many of the channels having a significant evoked response, the early post-stimulus portion of the AERP was found to have a clear characteristic oscillation with a period around 80 ms (12.5 Hz). In agreement, previous results on this data set (Bressler, Ding, & Liang, 1999) highlighted the existence of peaks near 12 Hz in power and coherence spectra for many cortical sites, especially during the early stages of the evoked response (up to approximately 200 ms after stimulus onset). The software for the AMVAR algorithm was coded by the author in MATLAB, and a standalone C executable was then generated with the MATLAB C-Compiler.

3.1.4 Absolute Phase Histograms

In order to compare predictions of the SPN and VSPN models regarding the distribution of phase, spectral phase estimates from Fourier transformed single trial residual time series were computed. In AMVAR spectral estimation, absolute phase information for single trials is lost and cannot be recovered. Alternative methods for calculating the instantaneous phase, such as the Hilbert transform and complex

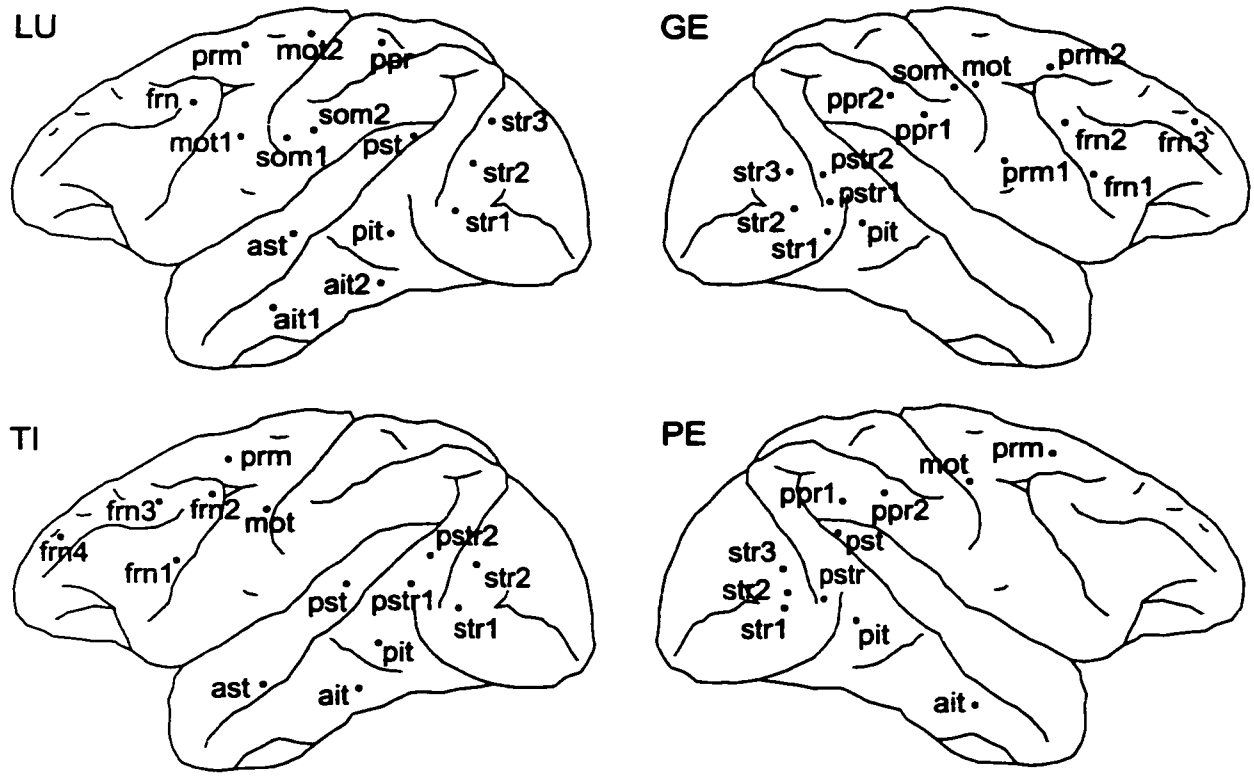

Figure 3-1: Sites for the transcortical electrodes placement. Labels: striate (str); prestriate (pst); posterior inferior temporal (pit); anterior inferior temporal (ait); posterior superior temporal (pst); anterior superior temporal (ast); posterior parietal (ppr); somatosensory (som); motor (mot); premotor (prm); frontal (frn). Recordings in LU and TI are from left hemisphere.

demodulation, were considered inadequate mainly because they would require narrow band-pass filtering of the data, and consequently high order filters and very long time segments of data. We approached the phase estimation problem in terms of optimal fitting in the least square sense. It is known that the Fourier transform of a time series at a specific frequency represents an optimal fit of the time series by sine and cosine functions at that frequency (Chatfield, 1995). For this reason, the FFT was used to obtain single trial phase estimates for constructing the phase distribution. The choice of window length for the FFT was constrained by three factors: a) the use of short windows to track the time evolution of the phase distributions from pre-stimulus to post-stimulus periods; b) the FFT constraint on the window length being a power of 2 (the alternative of zero padding introduces artifacts in the computation of phase distributions); and c) the choice of the 12 Hz Fourier component as an indicator of the main characteristic frequency of the AERPs . These constraints resulted in the following procedure. FFTs were performed on a time window (80 ms long) moving one data point at a time, from 50 ms before the stimulus onset to 350 ms after. The phase estimates of the 12 Hz (12.5 Hz being the actual frequency since we are using a 80 ms window) components for a window starting at time t, from the whole ensemble of single trials, were then collected to build the phase histogram at time t for a specific channel and subject. Each histogram's bin width was set to $\pi/50$ radians.

3.1.5 Testing for Departure from Uniform Phase Distribution

The SPN model predicts a uniform phase distribution. To assess whether a given absolute phase distribution departed from the uniform distribution, the modified Kuiper test (Fisher, 1995) for circular data was employed. In this procedure, for a specific time window, the linear order statistics $x_1 = \phi_1 / 2\pi, \ldots, x_r = \phi_r / 2\pi$ were first calculated, where ϕ_r was the phase of the 12 Hz Fourier component for trial r. Second, we obtained from the ensemble of R trials the statistics:

$$D_R^+ = \max(\frac{1}{R} - x_1, \frac{2}{R} - x_2, \ldots, 1 - x_r)$$
$$D_R^- = \max(x_1, x_2 - \frac{1}{R}, x_3 - \frac{2}{R}, \ldots, x_r - \frac{R-1}{R}) \quad (3.1)$$

$$V_R = D_R^+ + D_R^-. \quad (3.2)$$

Third, we computed the final statistic $V = V_R(R^{1/2} + 0.155 + 0.24 / R^{1/2})$. V was compared to critical values (V_c) given by the Kuiper modified statistics. At the significance level of $p < 0.01$, $V_c(0.01) = 2.0$. Since the phase distributions were computed as functions of time with the sliding window approach, the Kuiper V was also calculated as a time function for each channel from each monkey.

3.1.6 Cross-Correlation Time Function

The cross-correlation at time t between the residual time series from a channel pair $z_1(t)$ and $z_2(t)$ was computed using

$$C_{z_1 z_2}(\tau,t)=\frac{\sum_{r=1}^{R}[z_{1r}(t)-<z_{1r}(t)>_r][z_{2r}(t-\tau)-<z_{2r}(t-\tau)>_r]}{\sqrt{\sum_{r=1}^{R}[z_{1r}(t)-<z_{1r}(t)>_r]^2\sum_{r=1}^{R}[z_{2r}(t-\tau)-<z_{2r}(t-\tau)>_r]^2}}, \quad (3.3)$$

where τ is the time lag and $<z_{1r}(t)>_r$ denotes the trial ensemble average at time t. The cross-correlation was smoothed by computing it on the low-pass filtered (-3db at 22 Hz) and pre-processed residual time series. Zero-phase forward and reverse digital filtering was employed to prevent phase distortion (FIR2 and FILTFILT functions from MATLAB, version 6, Natick, MA).

3.2 Single Channel Quantities: AERP, Variance and Power

Predictions derived from the nonlinear gain modulation and from the VSPN model in Chapter 2 were tested using the macaque LFP data. Figure 3-2 shows example time functions of AERP, variance, and power at 12 Hz (see Methods for the choice of this specific frequency) from selected channels for each of the four monkeys. The AERP functions often exhibited several extrema. The variance function followed the expected temporal profile from the VSPN model with amplitude variability as the main contributor: the peak times roughly coincided with the times of one of the first two extrema of the AERP. The power at 12 Hz computed from the AMVAR model also followed the same predicted time course but in most cases with less temporal resolution, since its computation required a time window of 50 ms duration, which smeared the temporal structure.

We note that not every channel showed clear AERPs and peaks in the variance functions. For those channels that exhibited clear peaks, we examined the predicted correspondence between the time of the largest variance peak and the time of the largest AERP extremum. The scatter plot in Figure 3-3 combines the data from 30 channels from the four monkeys. The plot displays a general monotonic relation with points clustering around a line having slope of 0.74 (the expected slope from the VSPN model for the case of amplitude variability is a slope close to 1). The computed correlation coefficient was R = 0.87 ($R^2 = 0.76$). The same analysis for each individual monkey yielded the following results: GE, R = 0.94, including 6 channels in the striate, pre-striate, and parietal cortices; LU: R=0.72, including 8 channels in the striate, pre-striate, inferior temporal, and superior temporal cortices; TI: R=0.89, 9 channels in the striate, pre-striate, temporal, and frontal cortices; and PE: R=0.97, 7 channels in the striate, pre-striate, inferior temporal and pre-motor.

Typical examples of power modulation in different frequency bands are shown in Figure 3-4. Commonly, during the event-relate response (as evident by the AERP, power increases around 12 Hz are accompanied by decrease in power in higher frequency bands. A particularly striking feature is the transient power modulation in the parietal sites. An example is provided in Figure 3-5. The single trial time series show pre-stimulus activity concentrated around 20 Hz, which is "turned off" just after sensory activation, returning after the event-related response is over. The power spectrum clearly reflects this phenomenon.

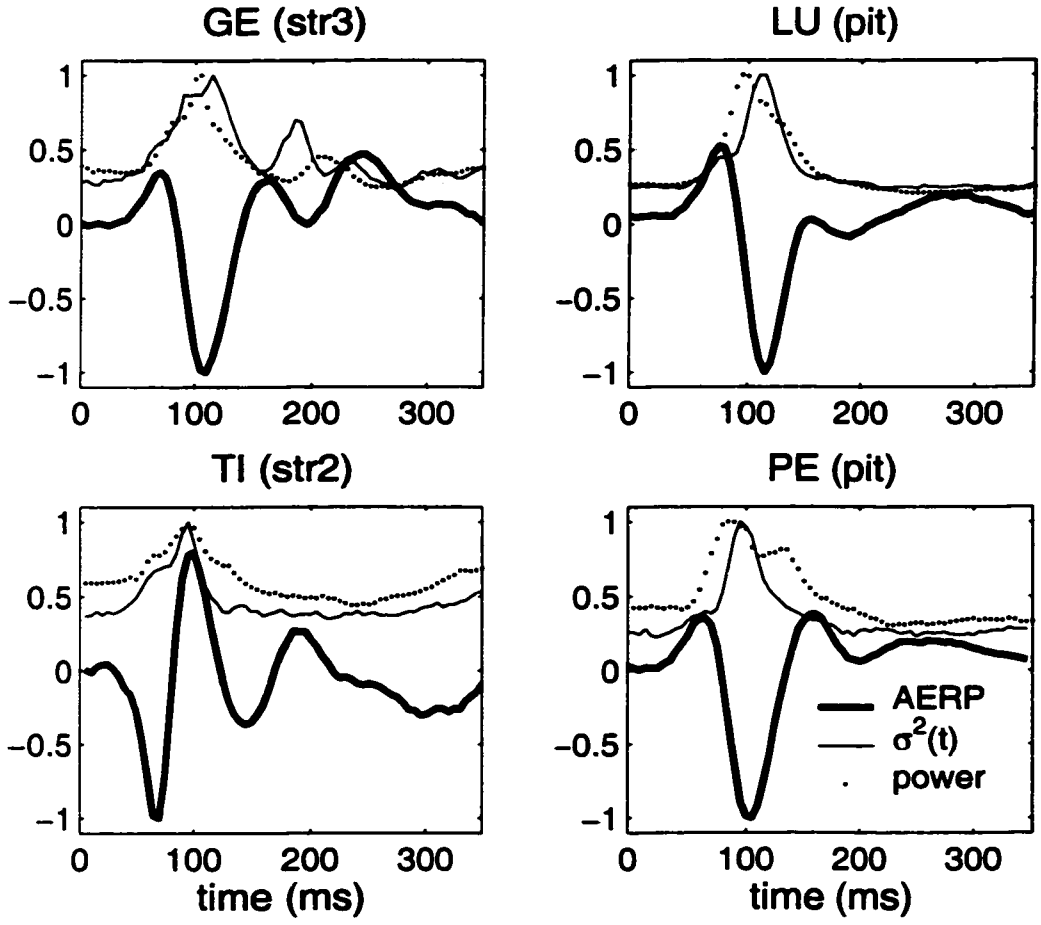

Figure 3-2: Example time functions of ensemble mean (AERP), variance (σ^2 (*t*)) and power at 12 Hz. To facilitate the comparison between the shapes, all quantities were normalized by their own maximal amplitude. Stimulus onset is at time 0 ms.

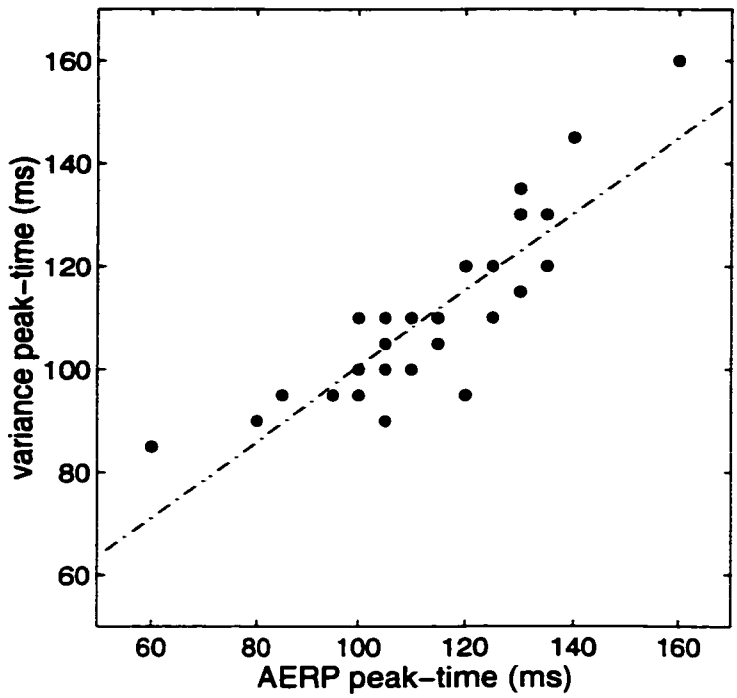

Figure 3-3: Time of the largest peak in the variance function plotted against the time of the largest AERP extremum. Results from 30 channels from the four monkeys are shown. The dashed line is the least square linear fit (slope equal to 0.74) relating the peak-times of the AERP and variance time function. The VSPN model predicts a slope close to 1 for the case where the amplitude is the main source of trial-to-trial variability.

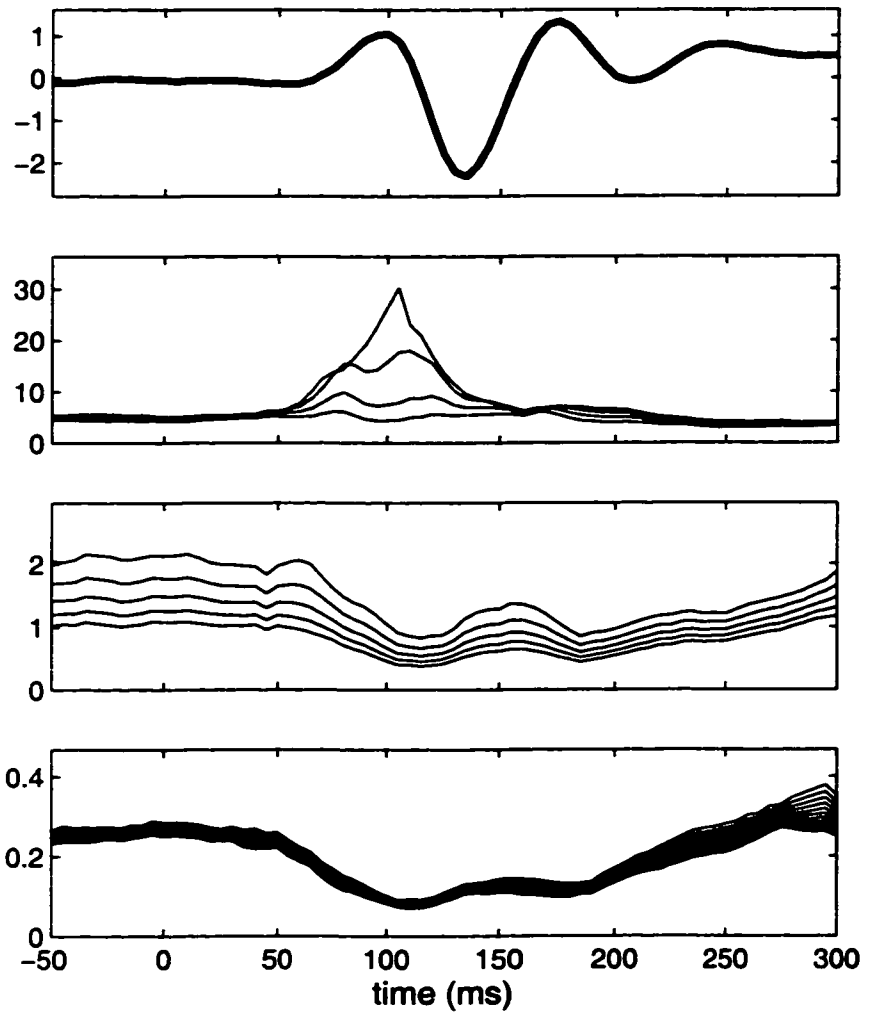

Figure 3-4: Example of an AERP function and power functions from a pre-striate (pst2) site from the monkey GE. The top plot shows the AERP. In the three bottom plots, the power spectral density function is shown. Each line represents the power at a specific frequency computed from a short (50 ms) moving window (see Methods). In the second plot (from top to bottom): 9 to 12 Hz (peak in power at 12Hz); third plot: 20 to 24 Hz; and fourth plot: 40 to 50 Hz. In the last two plots, power decreases with increasing frequency.

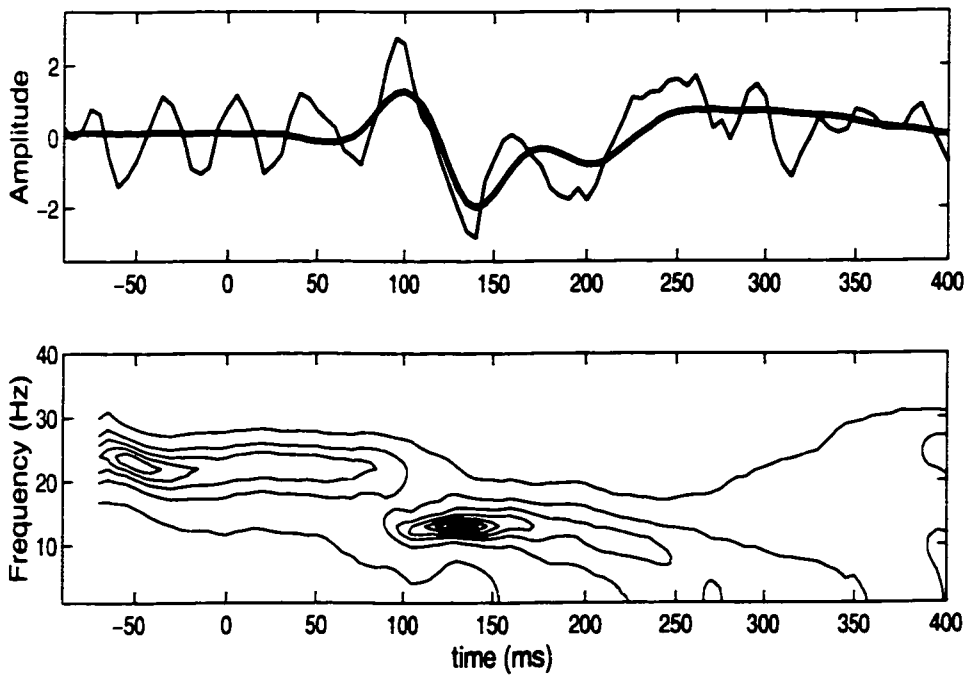

Figure 3-5: Event-related modulation of ongoing ~ 20 Hz oscillation. Top plot shows a single-trial example of abrupt perturbation of ongoing ~ 20 Hz oscillation at a parietal site (ppr1 monkey GE) with the advent of the stimulus-evoked response (the AERP appears in the thicker line). The single-trial amplitude has been subtracted by its own mean and normalized by its standard deviation. The bottom panel shows the corresponding contour plot of the power spectral density time function. Peaks in power are represented by the centers. Note the shift of the peak from ~ 22 Hz to ~12 Hz. The same phenomenon appears in recordings form somatosensory and motor cortices, and it is also accompanied by modulation in coherence. During the pre-stimulus period, there is significant coherence among all the parietal, somatosensory and motor cortices, suggesting that they form a preparation network. After the initial states of sensory stimulation, this beta coherence pattern is then replaced by coherence around 12 Hz, as mentioned before, simultaneously to the changes in power (see Figure 3-12).

Bimodal Phase Distributions

When amplitude variability is the main factor in the trial-to-trial variability of the evoked response, the VSPN model predicts a bimodal phase distribution for the residuals. This was tested by computing phase distributions of the 12 Hz component over the length of the trial. The Kuiper V statistic was computed for each phase distribution. Phase distributions having a Kuiper V value greater than $V_c = 2.0$ were judged to be significantly different ($p<0.01$) from the uniform distribution.

Figure 3-6 shows time functions of the Kuiper V statistic for selected channels from each monkey. As predicted by the VSPN model, the departure from uniform phase distribution is only observed following stimulus onset, during the stimulus evoked response. We found that all channels for all four monkeys had uniform phase distributions in the pre-stimulus period. By contrast, most channels in each monkey showed departure from the uniform distribution during stimulus processing. Notice that the maximum of the Kuiper V function tends to occur near 100 ms, which is about the time when maximal variance is observed (Figure 3-2). Table 1 lists the maximum Kuiper V computed during the post-stimulus period (0-200 ms) for all the channels and subjects.

Since the Kuiper statistic only tells whether or not a distribution is uniform, the phase distributions that significantly departed from uniformity were visually inspected to further determine whether or not they were bimodal. Figure 3-7 shows a representative phase distribution from each monkey, demonstrating clear-cut bimodality. For all the channels where we observed departure from the uniform distribution, the phase

distributions were ascertained to have bimodality. By comparison, an example of a uniform distribution from a pre-stimulus time window is presented in Figure 3-8.

Table 3.1. Maximum Kuiper V during post-stimulus period: 0-200 ms. (Critical $V_c = 2.0$, $p < 0.01$). See Figure 3-1 for the site label references.

Site	GE	LU	TI	PE
str1	**2.7**	**4.2**	**5.5**	**2.4**
str2	**3.5**	**6.1**	**4.7**	**2.7**
str3	**4.2**	**6.0**		**4.3**
pstr				**3.3**
pstr1	**4.2**		**4.0**	
pstr2	**1.9**		**2.6**	
pit	**3.7**	**6.1**	**6.7**	**4.6**
ait			**4.2**	**4.7**
ait1		**2.2**		
ait2		**3.7**		
pst		**4.1**	**1.4**	**1.5**
ast		**2.2**	**2.4**	
ppr		**1.8**		
ppr1	**3.7**			**1.8**
ppr2	**3.9**			**1.6**
som	**1.8**			
som1		**1.8**		
som2		**2.1**		
mot	**2.3**		**2.1**	**3.3**
mot1		**2.1**		
mot2		**1.2**		
prm		**3.5**	**3.4**	**1.7**
prm1	**2.7**			
prm2	**2.4**			
frn		**2.8**		
frn1	**2.8**		**2.4**	
frn2	**1.8**		**3.5**	
frn3	**2.2**		**2.7**	
frn4			**2.8**	

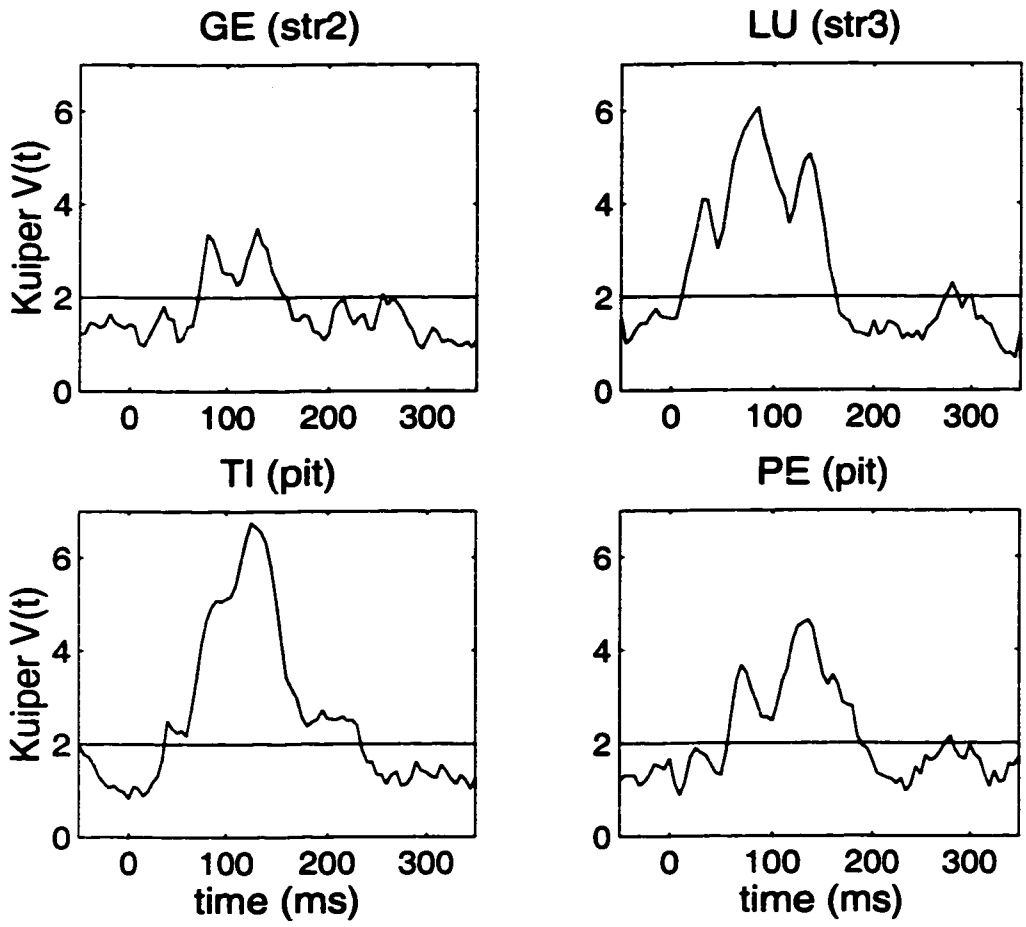

Figure 3-6: Four representative Kuiper V time functions. Notice that, during the pre-stimulus period, the absolute phase distributions of the 12 Hz Fourier components are not significantly different from uniform. During the evoked response, the distributions depart significantly from the uniform distribution. The horizontal line represents the critical level $V_c = 2.0$. Stimulus onset is at time 0 ms. The phase of the 12 Hz component was obtained by applying FFT on data segments from a moving window centered at time t.

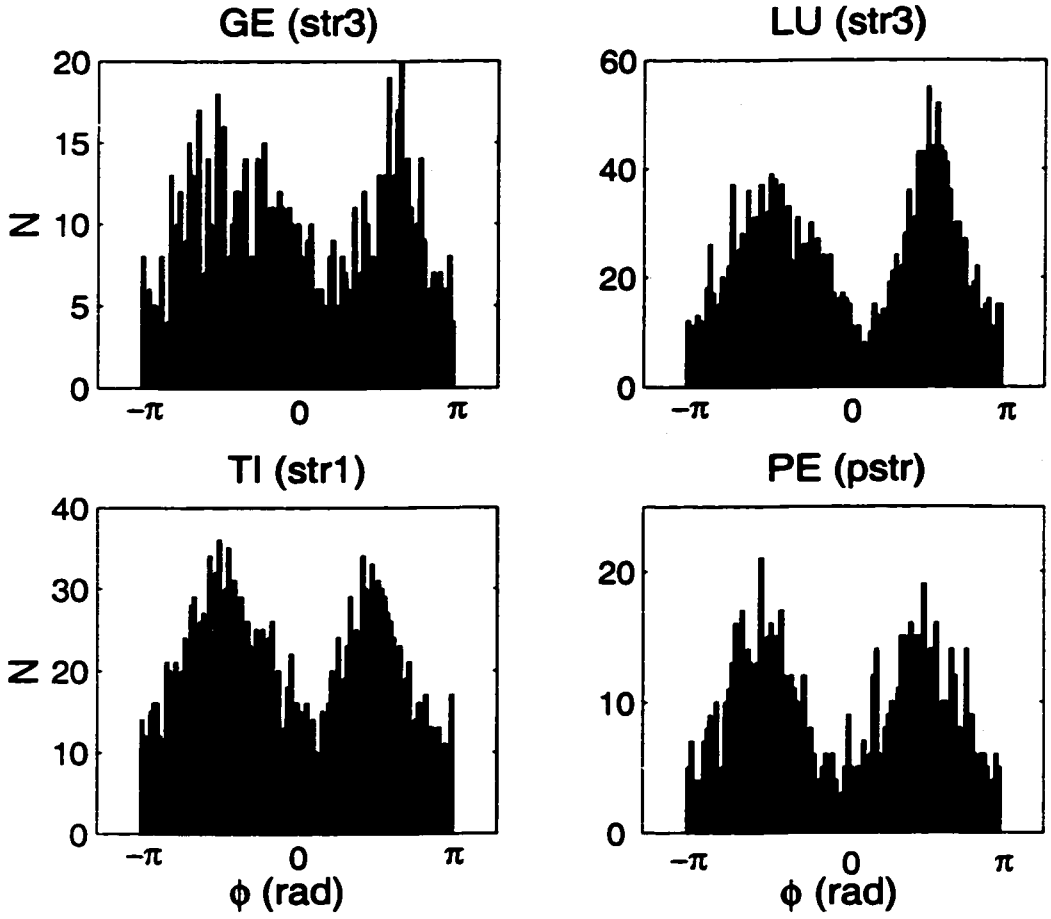

Figure 3-7: Bimodal phase. Four representative examples of absolute phase distributions of the Fourier component at 12 Hz (see Methods), computed from a time window centered at the times 95 ms (GE), 80 ms (LU), 80 ms (TI) and 75 ms (PE) during the evoked response period. The modes of the distribution are separated by approximately π radians, as expected from the VSPN model having amplitude variability as the main factor in the trial-to-trial variability of the evoked responses. The dispersion around the modes depends on the signal-to-noise ratio of the stimulus phase-locked component in the residuals and also on the latency variability of the evoked response.

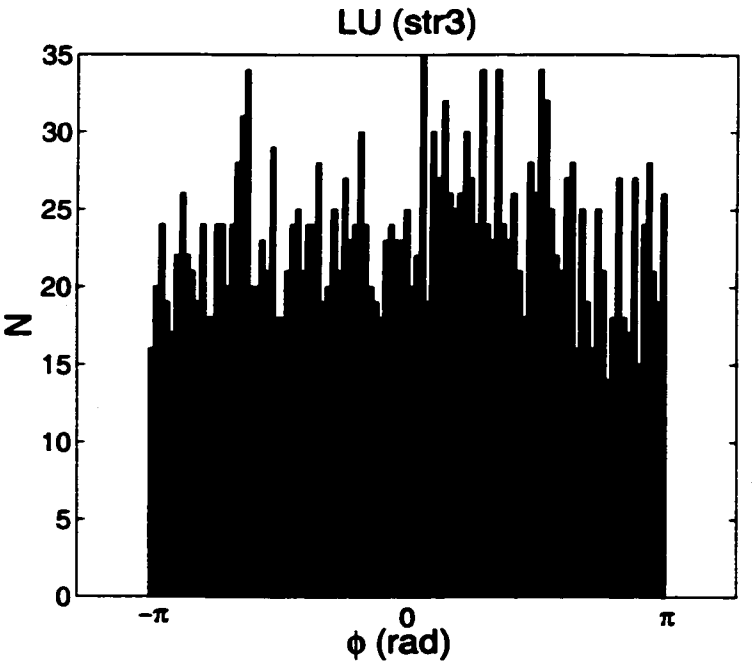

Figure 3-8: Uniform phase. Example of a uniformly distributed absolute phase distribution of the Fourier component at 12 Hz (the time window is centered at 10 ms after stimulus onset).

3.3 Cross-Correlation and Spectral Coherence

Modulations in cross-correlation time functions were observed to follow similar time courses as the signal-to-noise ratio at a variety of cortical sites, including striate, prestriate, inferior temporal, parietal, motor and prefrontal. Examples are shown in Figure 3-9. The oscillatory nature of the cross-correlation time function predicted by the VSPN model, and the temporal relation of its peaks with peaks in variance, are evident. The period of the oscillations in the cross-correlation functions agrees roughly with that in the variance functions, and is half the period of the main AERP oscillatory component, as expected from the VSPN model.

To estimate the prevalence of this effect, we needed to examine the time relation between the time peaks in variance and cross-correlation for the whole data set. A problem that complicated this analysis was that the AERPs and variance functions from different channels often did not peak at the same time. This discrepancy was perhaps due to temporal delays in direct or indirect transmission between the channel pair, or to the difference in the arrival times of a common input signal to both channels. To circumvent this difficulty, we chose the time lag τ in the cross-correlation function as the separation time between the largest variance peaks from each channel in the concerned pair. When the main characteristic frequency of the AERPs from two channels are similar and amplitude variation is the main factor for the temporal modulation of variance, then the VSPN model predicts that the time of the largest peak in the cross-correlation function, lagged by τ, should coincide roughly with the time of the earliest large variance peak of the two channels. This prediction was tested in a total of 144 channel pairs from the four

monkeys, selected according to the following criteria: (1) existence of clear peaks in the AERPs and variance functions; (2) the existence of non-uniform phase distributions events during the evoked response period; and (3) exceeding a conservative cross-correlation threshold (|cross-correlation extremum| ≥ 0.2). Figure 3-10 shows the relation between cross-correlation and first variance peak times. Superimposed is the line (slope ~ 0.94) obtained from the least square linear fit for the relation between peak-times in variance and cross-correlation. The VSPN model predicts a slope close to 1 for the case where the trial-to-trial variability in amplitude is the main contributor factor to fluctuations in signal-to-noise ratio. The correlation coefficient for the data in Figure 3-10 is R = 0.87 (R^2 = 0.76), supporting the role of the time-dependent signal-to-noise effect on the temporal modulation of the cross-correlation function. The coefficients computed individually for each monkey were: LU: 44 channel pairs, R = 0.90; GE: 41 pairs, R = 0.92; TI: 43 pairs, R = 0.86; PE: 16 pairs, R = 0.94.

Similar time-dependent signal-to-noise ratio effects were also observed for coherence time functions. Two important characteristics were detected. First, a commonly observed effect was the existence of peaks around 12 Hz in the coherence time function during the post-stimulus period. Those peaks tended to coincide with 12 Hz peaks in the power spectrum time function of at least one of the two channels contributing to the coherence. As shown in the beginning of this section, power peak times were correlated with changes in the ensemble variance time function and AERP. Figures 3-11, 3-12, 3-13 and 3-14 illustrates the relation between power and coherence peaks at 12 Hz. Second, since the computation of the spectral quantities required time segments of about 50 ms or

longer, the coherence time function resembled a smoothed version of the cross-correlation time function (Figure 3-15). Thus, although modulation in the coherence time functions clearly occurred, the relation between the peaks in coherence and the extrema of the AERP or the peaks in variance, was not always as evident.

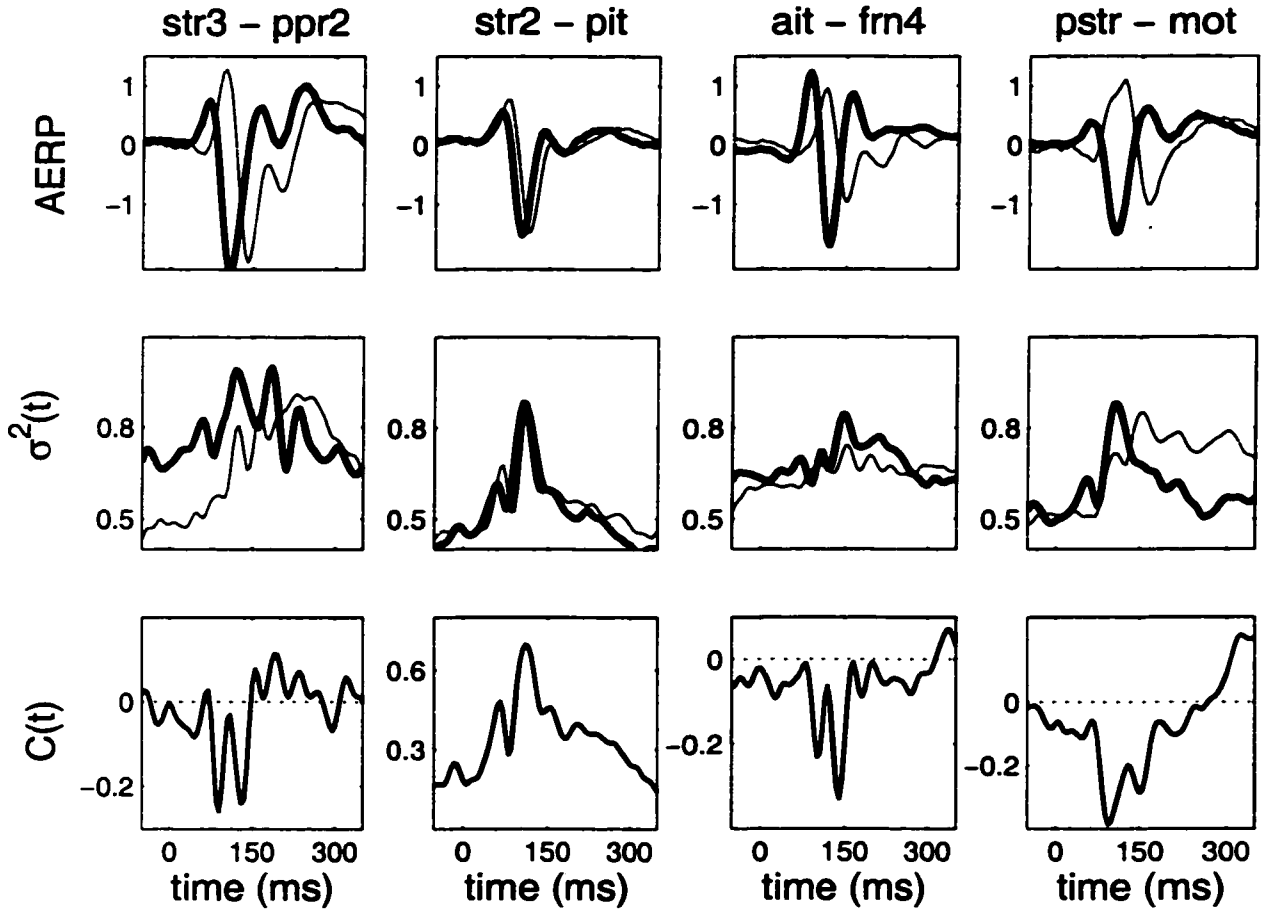

Figure 3-9: Oscillatory modulation of cross-correlation time functions. Illustrative examples of zero-lag point-by-point cross-correlation coefficients from four channel pairs are shown (bottom row) with the respective channels' AERPs (top row) and ensemble variances (center row), for four monkeys (from left to right: GE, LU, TI and PE). The variance was computed from low-pass filtered (-3 dB at 22 Hz), AERP-subtracted, and variance-normalized time series. Stimulus onset is at time 0 ms. The period of the oscillations in both variance and cross-correlation functions corresponds approximately to half the period of the main characteristic oscillation of the AERPs. Furthermore, peaks in

variance of at least one of the channel in the pair and extrema in cross-correlation tend to coincide in time.

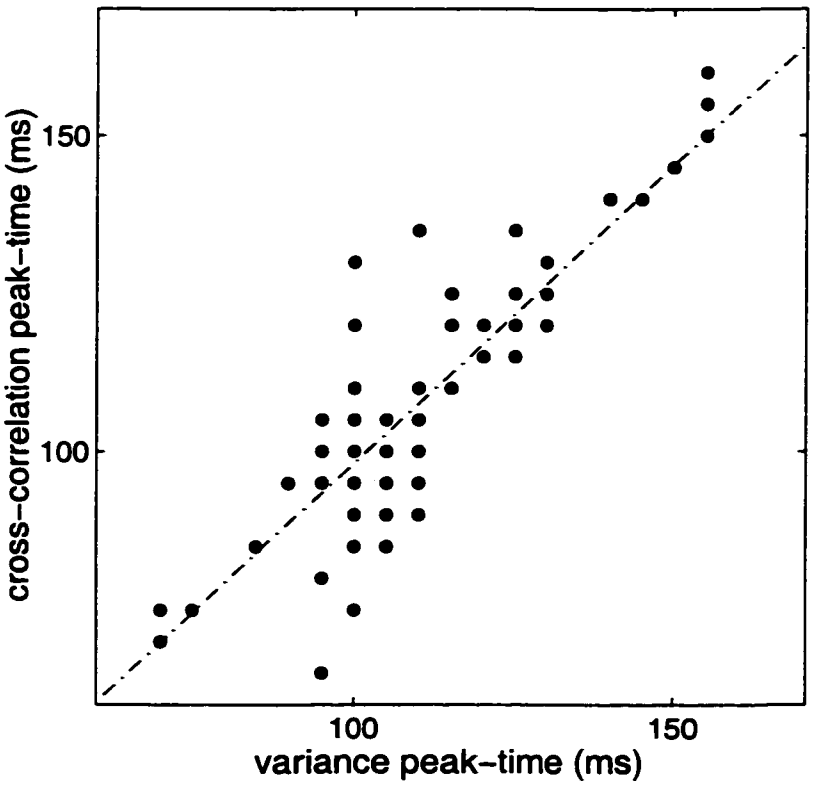

Figure 3-10: Peak-time of the lagged point by point cross-correlation time function versus the peak-time of the variance function for the channel with the earliest peak in variance in the pair (see Methods and text). Data from 144 channel pairs selected from the 4 monkeys are shown (see text). Because of the time resolution employed (sampling interval of 5 ms), many points are superimposed. The dashed line is given from the least square linear fit (slope equal to 0.94). The VSPN model predicts a slope close to 1 for the case where amplitude variation is the key factor for the temporal modulation in the signal-to-noise ratio.

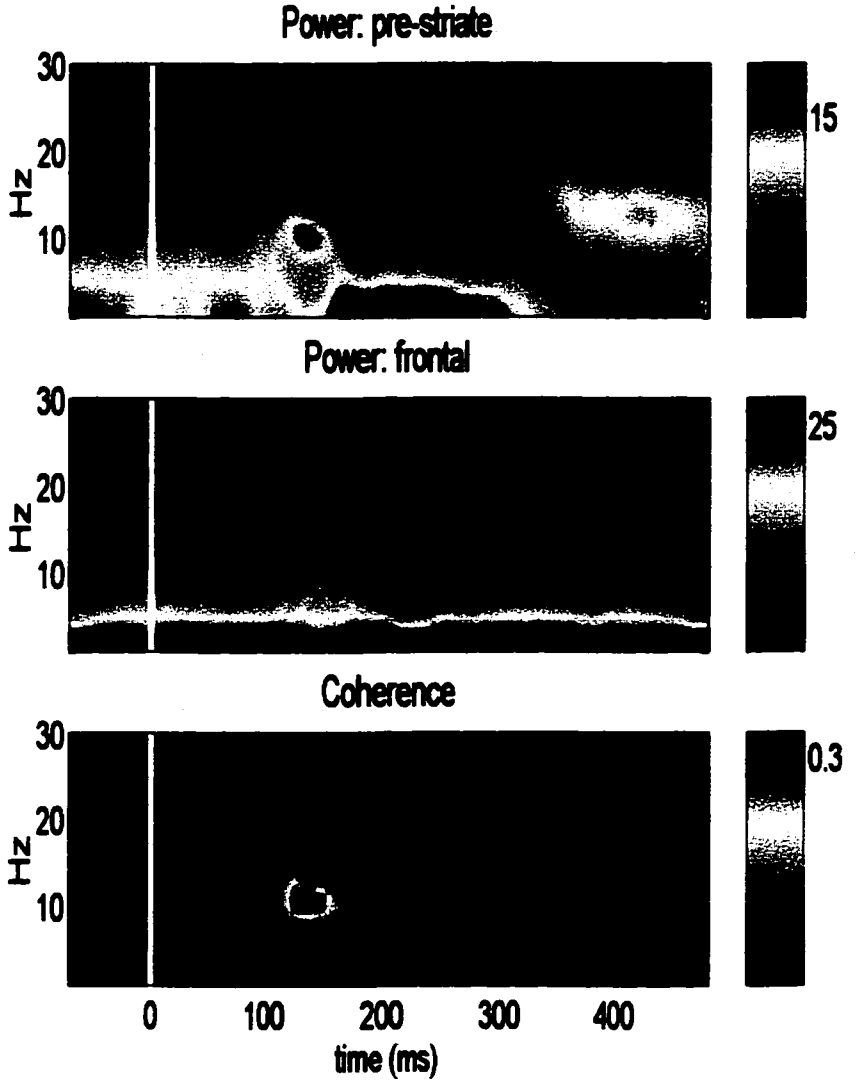

Figure 3-11: Event-related changes in power spectrum density and coherence time functions I. In both pre-striate (prst1) and frontal (frn1) channels (monkey GE), the event-related changes in power and coherence concentrate around the 12 Hz component. Also, peaks in power and coherence roughly coincide in time. The spectral quantities were derived from the AMVAR modeling of the residual time series (see Methods). Stimulus onset is at time 0 ms. Squared coherence values are used.

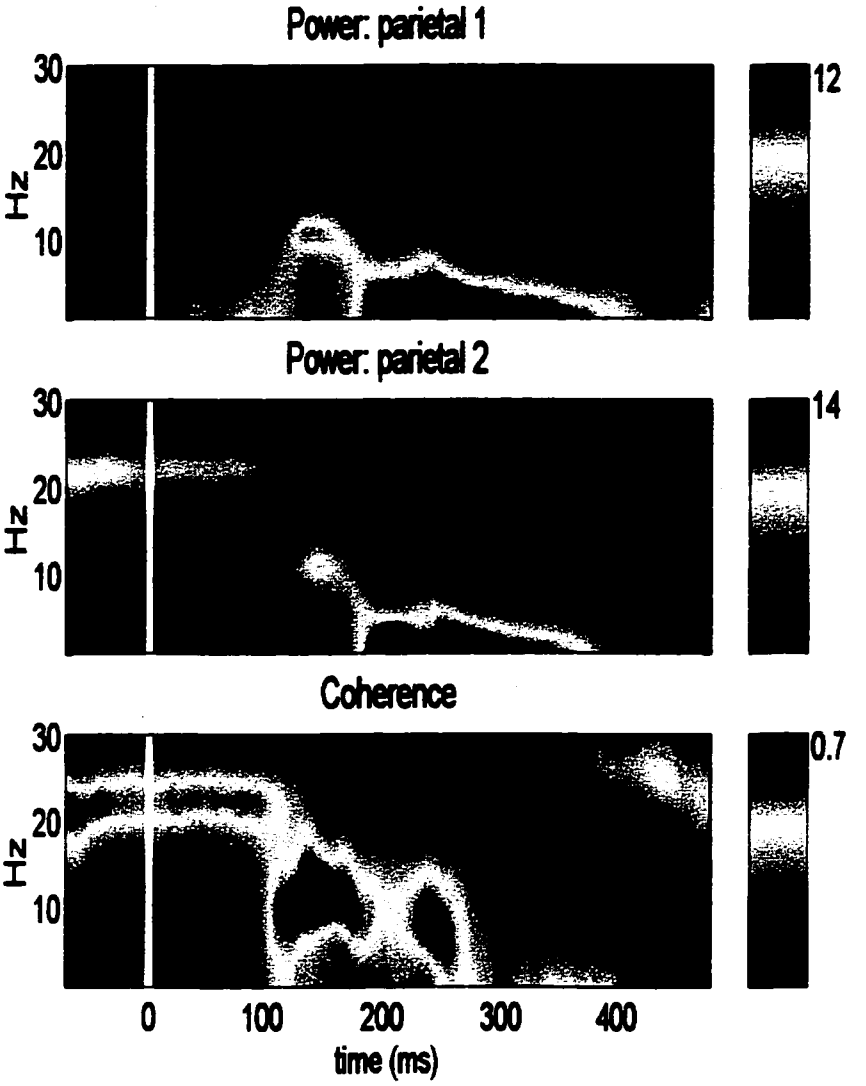

Figure 3-12: Event-related changes in power spectrum density and coherence time functions II. In both parietal channels (monkey GE, ppr1 and ppr2), the event-related changes in power and coherence concentrate around the 12 Hz component, with a simultaneous drastic decrease in the ~ 22Hz power and coherence. The same type of pattern appears among channel recordings from somatosensory and motor cortices. Peaks in power and coherence roughly coincide in time. The spectral quantities were derived from the AMVAR modeling of the residual time series (see Methods). Stimulus onset is at time 0 ms. Squared coherence values are used.

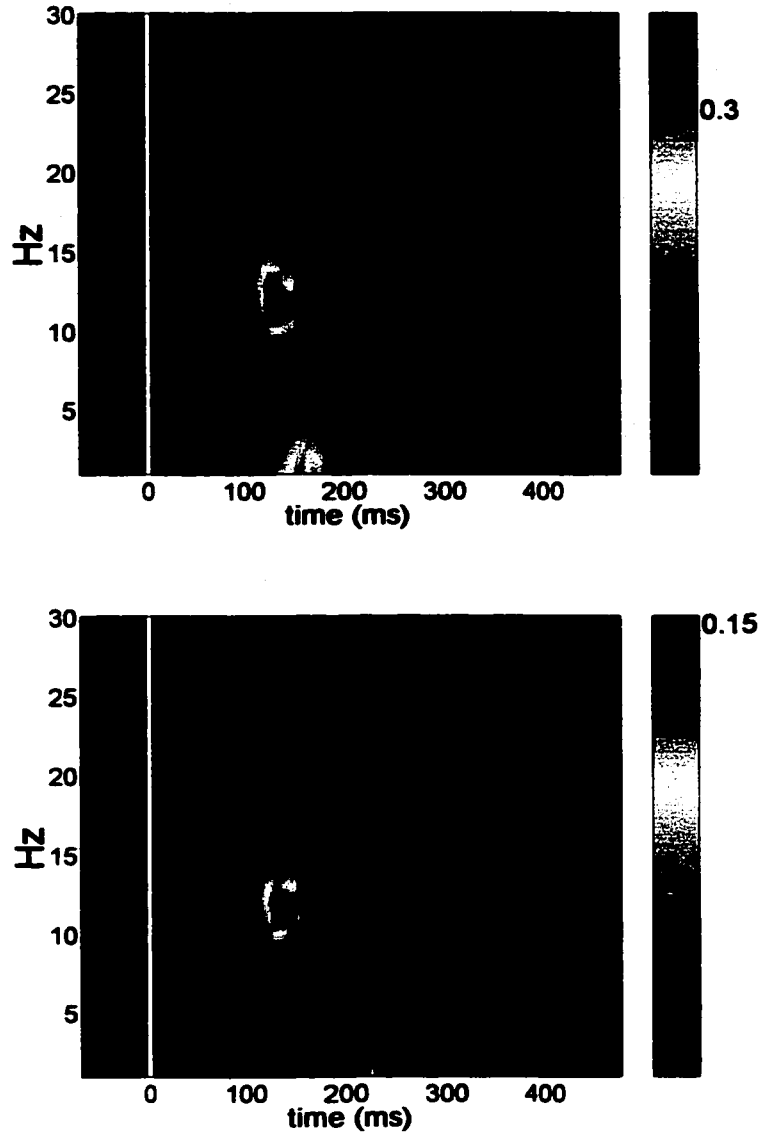

Figure 3-13: Event-related changes in coherence spectrum time functions. Another examples (monkey GE) for striate (str3) and pre-striate (pstr1), top plot, and striate (str3) and motor (mot), bottom plot, where the same transient in 12 Hz coherence is observed. The spectral quantities were derived from the AMVAR modeling of the residual time series (see Methods). Stimulus onset is at time 0 ms. Squared coherence values are used.

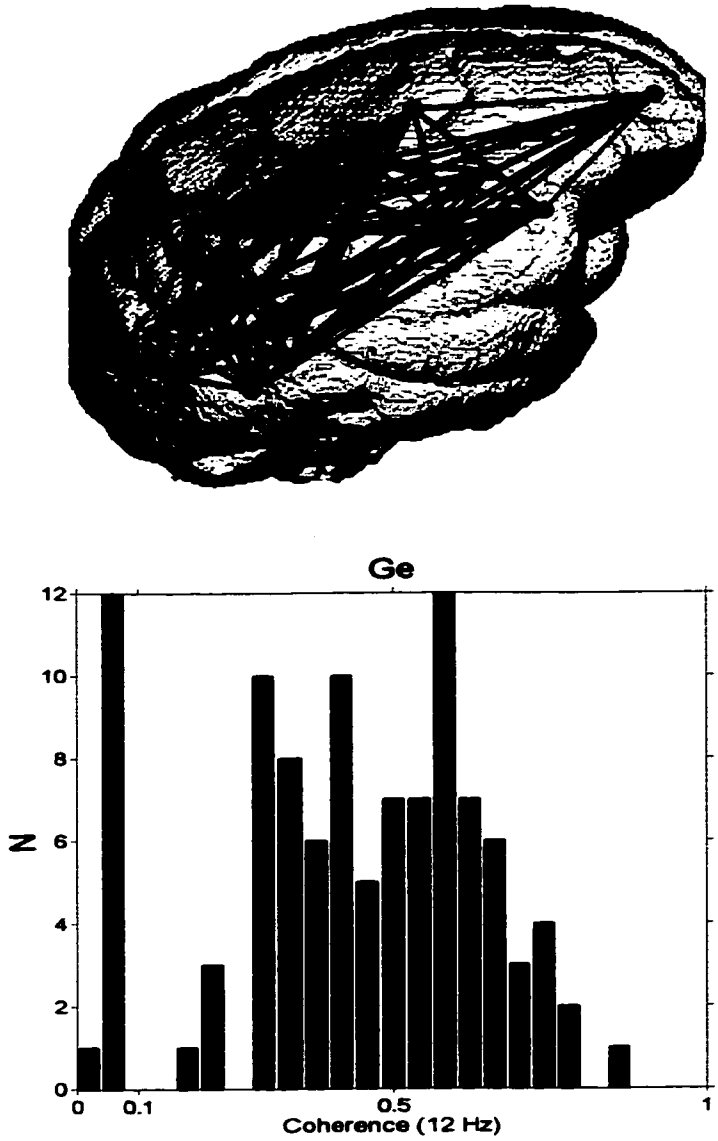

Figure 3-14: Event related transient increase in 12Hz coherence in monkey GE: summary. Around 120 ms there is a large widespread transient increase in the 12Hz coherence among most of the channel pairs as evidenced by the colored lines linking channel pairs with significant coherence (top plot). This widespread event is accompanied by transient increases in power in the same frequency as well (not shown). A histogram for the maximum coherence values around 120 ms is shown in the bottom plot, computed over 155 channel pairs. Significance level for the coherence values were established through bootstrap with resampling. In general, our results indicate that coherence values higher than 0.1 are statistically significant. Similar summaries for the other three monkeys are shown in Figure 4-8.

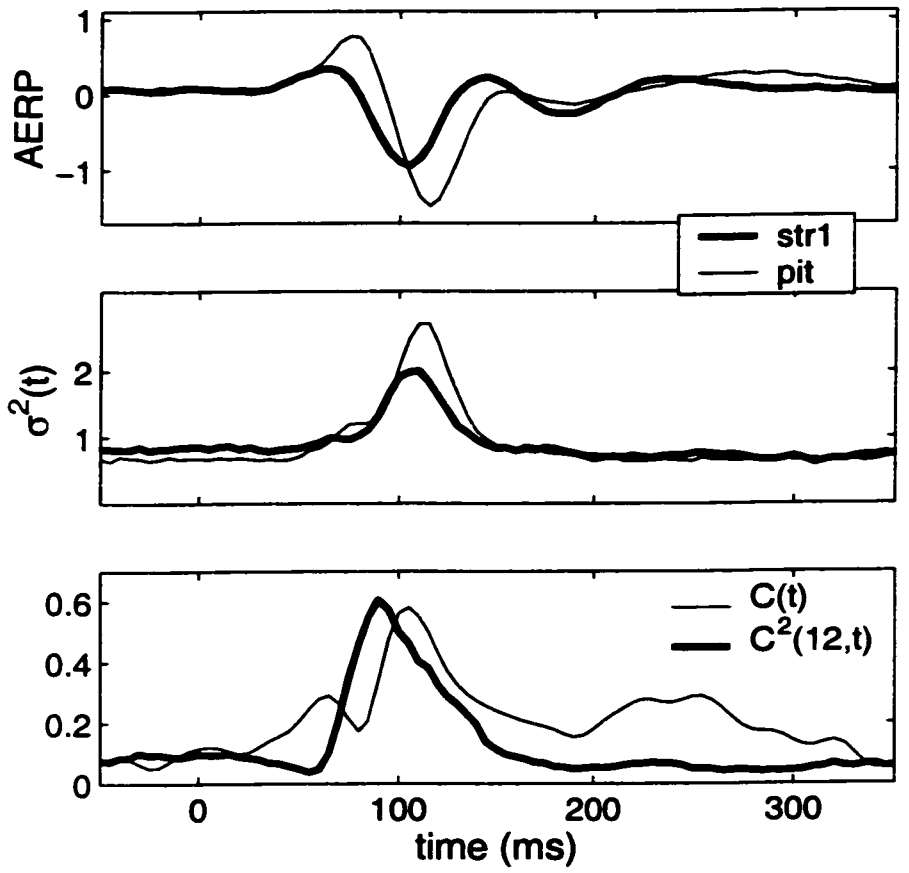

Figure 3-15: The time-dependent signal-to-noise ratio effect on the coherence time function. The shape of the zero-lag cross-correlation $C(t)$ is clearly related to the ensemble variance and AERPs. The squared coherence time function of the 12 Hz component $C^2(12,t)$ resembles a smoothed version of the cross-correlation time function, because of its computation within a time window. The plotted quantities refer to striate and posterior inferior temporal sites (monkey LU).

3.4 Discussion

Overall our results point to the predominance of the effect due to trial-to-trial variability of amplitude and latency on the statistical measures. The fact that bimodal histograms were observed constitute a strong indication that phase locked components exists in the residual time series as a result of amplitude variability (Figures 3-6 and 3-7). Thus the trial-to-trial variability is in place, and the remaining question is about the strength of its predicted effects. The modulation in power and variance and their relations to the AERP initially suggest the possibility of the gain modulation according to the proposed nonlinear model. However, more careful examination reveals this model's shortcomings. For instance, we would expect from the nonlinear model that an extrema of the AERP could coincide in time with peaks in variance and power, but not two opposed extrema as observed in some cases (Figure 3-2). This speaks more in favor of the amplitude variability. Further, contrary to the prediction that the gain modulation would be similar for a range of different frequencies (given the broadband characteristic of the pre-stimulus ongoing activity), we observe power increases in the alpha band, and decreases in the high frequency bands during the evoked response period (Figure 3-4). Another remarkable example of this abrupt decrease in high frequency power from the pre-stimulus to pos-stimulus period is the termination of the pre-stimulus narrow-band (~20Hz) oscillation in parietal sites (Figures 3-5, 3-12). This event is also associated with decreases in coherence in the same frequency among several sites in the parietal, somatosensory and motor cortices. This type of modulation in the beta and gamma bands could be the result of gain modulation. However, it seems to exhibit a more complex

nature than the one presented by the proposed nonlinear model. For this reasons, we discuss the experimental results more in the context of the predictions of the VSPN model, especially those related to the trial-to-trial variability in amplitude. The topic of event-related modulation of ongoing activity will be addressed again in more detail in Chapter 4.

3.4.1 Variability of Cortical Recordings

In the traditional SPN model, trial-to-trial variability is entirely attributed to independent ongoing activity, with which a repeatable (trial-to-trial stationary) evoked response is linearly combined. The assumptions of linear combination and trial-to-trial stationarity have often been questioned in the past: nonstationarity of the evoked response has been considered as another possible source of variability in EEG and LFP recordings, while multiplicative effects between ongoing and evoked activities have been treated as leading to transient modulation of both of their variances. Substantial experimental support for variability of the evoked response, including amplitude, latency and even waveform variability, has been provided in previous studies (Coppola, Tabor, & Buchsbaum, 1978; Horvath, 1969; Lange, Pratt, & Inbar, 1997; Mocks, Gasser, Tuan, & Kohler, 1987; Woody, 1967). The issue of whether the neural system behaves in a linear regime during the evoked response has been more controversial, with experimental evidence either supporting {Ref. Arieli, Sterkin, Grinvald, & Aertsen, 1996; Azouz & Gray, 1999} or contradicting the linearity assumption (Aertsen, Erb, & Palm, 1994;

Chawla, Lumer, & Friston, 2000; Kisley & Gerstein, 1999; Vijn, van Dijk, & Spekreijse, 1991).

As a first refinement of the SPN model, the VSPN model assumes that the stimulus-evoked response has a stereotyped waveform but that the amplitude and latency of this waveform can vary from trial to trial. A number of possible factors may contribute to this trial-to-trial variability. For example, slow changes of global brain state involving arousal and attention could lead to fluctuations in the excitability of cortical and subcortical neuronal populations, thereby causing variability in the magnitude and timing of the evoked response (Brody, 1998; Mangun & Hillyard, 1991), without significantly affecting the shape of the response. Also, in the context of visual evoked responses, eye movements could change the location of retinal stimulation and consequently the location of cortical excitation in retinotopically organized areas (Gur, Beylin, & Snodderly, 1997), contributing to variability of the recorded activity from fixed electrodes. Trial-to-trial variability could also be related to the previously observed dependence of single trial evoked response amplitude on the level of pre-stimulus ongoing activity (Basar, Gonder, & Ungan, 1976; Basar, Rahn, Demiralp, & Schurmann, 1998; Brandt & Jansen, 1991). Although all these sources of variability are commonly considered to be a violation of the stationarity assumed in ideally controlled experimental designs (Aertsen, Gerstein, Habib, & Palm, 1989), we argue that some of them, like state dependent neural excitability, constitute in practice an inevitable aspect of multi-trial experiments, even for situations where the recording session might last only a few minutes.

3.4.2 Observed Effects of the Trial-to-Trial Variability on Recorded Local Field Potentials

Analysis of the local field potential data was intended to verify the predicted effects of trial-to-trial variability of the evoked response in a very common type of experimental design. Several lines of evidence for amplitude variability were: (1) the predicted relationship between variance and AERP waveform; (2) the bimodal histograms for phase of the 12 Hz Fourier component computed on residual time series; In all four subjects, the majority of recorded channels showed phase distributions that, according to the Kuiper *V* statistics, departed from uniformity, supporting the existence of stimulus phase-locked signals in the residual time series. This effect was observed during the initial period of the evoked response (50-200 ms) in all recorded cortical regions, from the primary visual cortex to the motor and frontal cortices (Table 1 and Figure 3-5). Visual inspection of the non-uniform distributions showed them to be clearly bimodal in many cases (Figure 3-7). This implies that the signal-to-noise ratio was high and that the latency variability was not severe. Large latency variability would lead to broader modes in the phase distribution. In extreme cases, the distribution would look uniform. Low signal-to-noise ratios would result in the distributions being dominated by the 12 Hz component of the ongoing activity, which again would be expected to be uniform.

The VSPN model predicted that peaks in the ensemble variance should coincide in time with extrema of the AERP for cases where amplitude variability is the main contributing factor. Examination of the data revealed this to be approximately the case, especially for AERP extrema before 150 ms (Figures 3-2 and 3-3). Deviations from this

prediction possibly arose because of substantial latency variability. Analysis of the power spectrum time function revealed that the main increase in variance occurred for Fourier components around 12 Hz, which was the main characteristic frequency of the AERPs (Figures 3-2, 3-11 and 3-12).

Regarding interdependence measures, i.e. cross-correlation and spectral coherence time functions, the VSPN model's main implication was described in terms of a time-dependent signal-to-noise ratio effect. Two main scenarios were described in Chapter 2, Section 2.2.2. First, if the trial-to-trial amplitudes or latencies of the evoked responses recorded from two channels co-vary, the time-dependent signal-to-noise ratio effect implies that characteristic transient increases in the cross-correlation should be observed during the evoked response period. In the second scenario, the trial-to-trial amplitudes or latencies do not co-vary, but the pre-stimulus ongoing activity is correlated between the two channels. In this case, instead of a transient increase in the cross-correlation, as compared to the pre-stimulus period, a decrease should be observed. The results presented in this paper related exclusively to the case of increase in cross-correlation and coherence. As expected from the VSPN model, the event-related increase in cross-correlation time functions was accompanied by event-related increases in 12 Hz power and for many channel pairs, in event-related increase in coherence in the same frequency (Figures 3-11, 3-12 and 3-13).

The relation between variance peak times and cross-correlation peak times (Figure 3-10) was nearly linear with slope close to one, supporting the hypothesis of amplitude co-variation. Observed deviations may have originated again from latency variability or

from two other factors. A complete coincidence between the peak times would be observed if the AERPs of different channels were identical, which was not the case (see Chapter 2, Section 2.2.2 and Figure 2-8). In the scatter plot (Figure 3-10), many channel pairs were grouped together. Some of these channels may have had their statistical interdependence arising from higher order interactions (e.g., indirect interaction through one or more cortical areas). As in the hypothetical example of two parietal channels given above, even though a common striate input might be the main factor contributing to the modulation of their interdependence, it might only have a minor influence on the characteristic shape of their ensemble variances. In the spectral domain, a much more consistent time coincidence was observed between peaks in 12 Hz coherence and 12 Hz power (examples were provided in Figures 3-11, 3-12 and 3-13).

3.4.3 Separation of Stimulus Phase-Locked and Non-Phase-Locked Components in Cortical Recordings

Event-related recordings commonly contain evoked and/or induced components (Freeman, 1994; Pfurtscheller & Lopes da Silva, 1999) in addition to unrelated background activity. The evoked component is usually considered to result from phase-locked responses of cortical populations to thalamic afferents, while the induced component is time-locked but not phase-locked to the stimulus, and results from transient changes in neural parameters controlling interactions among cortical populations. The evoked component tends to have a low frequency profile that can exhibit phase-locking properties in the presence of latency variability. On the other hand, induced components

tend to have higher frequencies (e.g. in beta or gamma ranges) and need more careful analysis to discover. In the present work, induced activity and background activity are not differentiated and are collectively called the ongoing noise component.

In summary, the observed event-related transient increases in variance, power, correlation and coherence have a contribution from the SNR effect. What is left to ascertain is the strength of this contribution. The observed effects could be originating from combined nonlinear gain modulation (of a more complex nature than the one initially proposed) and nonstationarities of the event phase-locked components. To clarify this issue, a separation of the remnant phase-locked components in the residual time series and the ongoing activity will be attempt in the next Chapter where a estimation framework based on Bayesian inference is proposed.

Chapter 4

Bayesian Analysis of Single Trial Cortical Event-Related Components

A common paradigm for the study of the relations between brain activity and behavior is the recordings of event-related potentials over repeated presentations of a sensory stimulus or task demand. The obtained recordings are commonly understood as the linear combination of event-related signals relatively phase-locked to the event onset, and ongoing activity. The ongoing activity includes event unrelated activity and possibly also event-induced components that are not phase-locked but time-locked to the event onset. Based on the classic signal-plus-noise model, the most common application of this paradigm has been to capture the event-related potential by averaging across the ensemble of trials, resulting in the Averaged Event-Related Potential (AERP). A weakness of this

application is its assumption that the true event-related potential is invariant across trials with respect to its waveform, amplitude and latency. Extensive evidence has led to more realistic assumptions where trial-to-trial amplitude and latency variability are allowed. The existence of this trial-to-trial variability implies that the ensemble mean is no longer the optimal estimation of the event-related potential. Further, the separation between these components and the ongoing activity is fundamental for the interpretation of event-related modulation of statistical interdependence in multi-channel data as seen in the previous Chapters. Also, the estimation of single trial amplitudes and latencies of event-related potentials provides essential information for the understanding of information processing stages in the cortex (Nowak & Bullier, 1997).

Several attempts to provide estimates of single trial event-related response following maximum likelihood approaches have been presented before, e.g.: single trial ICA (Jung, et al., 1999), and latency and amplitude of single trial evoked responses (Jaskowski & Verleger, 1999; Lange, Pratt, & Inbar, 1997; Pham, Mocks, Kohler, & Gasser, 1987; Woody, 1967). Here we provide a more comprehensive formulation.

4.1 Generative Model and Single Trial Parameter Estimators

We start from a multivariate linear generative model (Truccolo, Knuth, Bressler, & Ding, in press) expressed as:

$$\begin{pmatrix} z_{1r}(t) \\ \vdots \\ z_{Mr}(t) \end{pmatrix} = \begin{pmatrix} B_{11} & \cdots & B_{1N} \\ \vdots & \cdots & \vdots \\ B_{M1} & \cdots & B_{MN} \end{pmatrix} \begin{pmatrix} S_{1r}(t) \\ \vdots \\ S_{Nr}(t) \end{pmatrix} + \begin{pmatrix} \eta_{1r}(t) \\ \vdots \\ \eta_{Mr}(t) \end{pmatrix}, \tag{4.1}$$

or in matrix form:

$$\mathbf{Z}_r(t) = \mathbf{B}\mathbf{S}_r(t) + \boldsymbol{\eta}_r(t), \tag{4.2}$$

where $z_{mr}(t)$ is the m^{th} channel recording at time t and for the r^{th} realization or trial. $S_{nr}(t)$ represents a particular realization of the n^{th} event or source activity at the trial r, B_{mn} are the entries of a mixing matrix representing instantaneous and stationary source-detector couplings, and $\eta_{mr}(t)$ refers to an unpredictable component or process whose density is specified by a set of parameters θ_η. The sources or components $S_{nr}(t)$ can be modeled in many diverse ways, as for example, ARMA processes, continuous or discrete time source waveforms, Taylor expansion of dynamical models, etc.

We approach the problem of solving for the components' amplitudes, latencies and waveforms, the parameters θ_η for the unpredictable components $\eta_{mr}(t)$ and the matrix **B**, from a Bayesian perspective. According to Bayes theorem, the posterior probability of model parameters M given the data D and prior information I, is given by:

$$p(M|D,I) = \frac{p(D|M,I)\ p(M|I)}{p(D)}, \tag{4.3}$$

where $p(D|M,I)$ is usually referred as the likelihood and $p(M|I)$ is the prior probability of the model.

For the particular generative model in question, the posterior probability of the model parameters given the data and prior information I is:

$$p(\{\mathbf{S}(t)\},\mathbf{B},\{\theta_\eta\}|\{\mathbf{Z}(t)\},I) = \frac{p(\{\mathbf{Z}(t)\}|\{\mathbf{S}(t)\},\mathbf{B},\{\theta_\eta\},I)\ p(\{\mathbf{S}(t)\},\mathbf{B},\{\theta_\eta\}|I)}{p(\{\mathbf{Z}(t)\})}, \tag{4.4}$$

where $\{\cdot\}$ represents the set of a specific parameter or data for the whole ensemble of channels, trials, components and time points, $p(\{\mathbf{Z}(t)\}|\{\mathbf{S}(t)\},\mathbf{B},\{\theta_\eta\},I)$ is the likelihood and $p(\{\mathbf{S}(t)\},\mathbf{B},\{\theta_\eta\}|I)$ is the prior probability for the model parameters. From a rigorous Bayesian perspective, the solution for the problem is the whole posterior probability. For high dimensional data sets, sampling directly from the posterior probability density is not feasible. The alternative is provided by Markov Chain Monte Carlo (MCMC) techniques (Chen, Shao, & Ibrahim, 2000), where a Markov Chain on the parameters is designed such that it converges to the posterior probability distribution in question. Once the system is in equilibrium, the means, modes, covariances and other higher moments for the parameters are computed from the sampled time series of the Markov chain.

Here, we consider a simpler approach by choosing the *Maximum a Posteriori* (MAP) solution, i.e., the set of parameters that maximize the posterior. In the context of Equation (4.3), the MAP solution for the model parameters M is

$$\hat{M} = \arg\max_M [p(D|M)p(M)] = \arg\max_M [\ln p(M) + \ln p(D|M)]. \qquad (4.6)$$

It is interesting to notice that this solution is related to the Minimum Description Length (MDL) principle seen in the Chapter 1, Section 1.2.2 when it was applied to the AR model selection (order). The essence of the MDL principle is to minimize the length (order) of a representation defined as:

$$|\mathcal{L}(M,D)| = |\mathcal{L}(M)| + |\mathcal{L}(D \text{ using } M)|, \qquad (4.7)$$

where $|\cdot|$ denotes the length of the description and $\mathcal{L}$ a description language. This can be intuitively understood in the framework of a communication model, where a stochastic

variable is being coded (Zemel, 1994). The sender takes as input a set of observations, and sends an encoded message that the receiver decodes and uses to produce the output. When the output of the receiver is identical to the input, then the message is called an information-preserving description of the input. The output of the sender is a message written in the description language $\mathcal{L}$ of the receiver. This message conveys both the model M, and the data D when encoded such that the model M can decode it; the term $|\mathcal{L}(D \text{ using } M)|$ is better understood as the errors or residuals, i.e., aspects of the data not predicted by the model. Explicitly, the MDL principle advocates:

$$\hat{M} = \arg\min_M |\mathcal{L}(M, D)|. \tag{4.8}$$

The link between the MAP solution, Equation (4.6), and the MDL solution, Equation (4.8), is provided by the Shanon Coding theorem (Cover & Thomas, 1991). Without attempting rigor and considering only the case where the true probability distributions (over discrete events with finite size representation) are known, the optimal code length for a stochastic variable X is given by:

$$|\mathcal{L}(x)| = -\ln[p(X = x)], \tag{4.9}$$

where $p(X = x)$ is the probability of the discrete event x. From this, Equation (4.8) can be rewritten as:

$$\hat{M} = \arg\min_M |\mathcal{L}(M, D)| = \arg\min_M [-\ln[p(M)] - \ln[p(D|M)], \tag{4.10}$$

which is equivalent to the MAP solution $\hat{M} = \arg\max_M [\ln p(M) + \ln p(D|M)]$.

It is known that the MAP solution of generative models similar to (4.2) gives origin, depending on the choice of prior distributions for the model parameters, to a set of

very important techniques. Examples include cluster analysis (K-means), PCA and Factor Analysis (Dayan & Abbott, 2001; Hinton & Ghahramani, 1997) (see chapter 10 in the last reference). Also, with certain special assumptions, as for example, noiseless recordings with invertible matrix **B**, and spatially stationary and trial invariant independent sources with non-Gaussian high kurtosis densities for their amplitudes, one can promptly derive Bell and Sejnowski's ICA algorithm (Knuth, 1997).

The above generative model is further specified in the context of event-related cortical local field potential (LFP) recordings. It improves upon the classic signal-plus-noise model (where linear superposition and trial-to-trial invariance are assumed) by allowing variability of the single trials amplitudes and latencies of an invariant waveform response. Consider the specific type of source activity representation:

$$\mathbf{S}_r(t) = \begin{pmatrix} a_{1r} s_1(t - \tau_{1r}) \\ \vdots \\ a_{Nr} s_N(t - \tau_{Nr}) \end{pmatrix}, \tag{4.11}$$

and re-express model Equation (4.2) in component form:

$$z_{mr}(t) = \sum_{n=1}^{N} B_{mn} a_{nr} s_n(t - \tau_{nr}) + \eta_{mr}(t), \tag{4.12}$$

where $s_n(t)$ is the n^{th} event-related response component with a trial-to-trial variable amplitude and latency given, respectively, by $a_{nr}(t)$ and $\tau_{nr}(t)$. The unpredictable component $\eta_{mr}(t)$, henceforth simply referred as ongoing activity, is now a zero mean activity including both the true ongoing activity and other uncertainties originating from measurement noise. The posterior probability can then be expressed as

$$p(\{s_n(t)\},\{a_{nr}\},\{\tau_{nr}\},\mathbf{B},\{\theta_\eta\}|\{z_{mr}(t)\},I) =$$

$$\frac{p(\{z_{mr}(t)\}|\{s_n(t)\},\{a_{nr}\},\{\tau_{nr}\},\mathbf{B},\{\theta_\eta\},I)\ p(\{s_n(t)\},\{a_{nr}\},\{\tau_{nr}\},\mathbf{B},\{\theta_\eta\}|I)}{p(\{z_{mr}(t)\})}$$

(4.13)

For this additive model, the likelihood $p(\{z_{mr}(t)\}|\{s_n(t)\},\{a_{nr}\},\{\tau_{nr}\},\mathbf{B},\{\theta_\eta\},I)$ turns out to be simply the probability model of the ongoing activity: $p(\eta(t)|\{\theta_\eta\},I)$. This can be shown by applying the transformation law for probabilities $p(x|I) = p(y|I)|dy/dx|$, or more easily using marginalization as in the following the example. Consider the additive model

$$y = x + \varepsilon. \tag{4.14}$$

From this, we write the conditional probability or likelihood

$$p(y|x,I) = \frac{p(x|y,I)p(y|I)}{p(x|I)} = \frac{p(y,x|I)}{p(x|I)}. \tag{4.15}$$

Notice that $p(y,x)$ can be derived applying the marginalization and multiplication rule of the probability calculus as

$$\begin{aligned} p(y,x|I) &= \int p(y,x,\varepsilon,I)d\varepsilon \\ &= \int p(y,x|\varepsilon,I)p(\varepsilon|I)d\varepsilon \\ &= \int p(y|x,\varepsilon,I)p(x|\varepsilon,I)p(\varepsilon|I)d\varepsilon. \end{aligned} \tag{4.16}$$

To enforce the condition on $p(y|x,\varepsilon,I)$ provided by the model $y = x + \varepsilon$, we assign a delta function

$$p(y|x,\varepsilon,I)=\delta(y-x-\varepsilon). \tag{4.17}$$

Thus, Equation (4.16) can be rewritten as

$$p(y,x|I)=\int \delta(y-x-\varepsilon)p(x|\varepsilon,I)p(\varepsilon|I)d\varepsilon. \tag{4.18}$$

or

$$p(y,x|I)=p(x|\varepsilon=y-x,I)p(\varepsilon=y-x|I). \tag{4.19}$$

Substituting Equation (4.19) into Equation (4.15), we obtain:

$$p(y|x,I)=\frac{p(x|\varepsilon=y-x,I)p(\varepsilon=y-x|I)}{p(x,I)}. \tag{4.20}$$

Finally, since x and ε are assumed to be independent, i.e., $p(x,\varepsilon|I)=p(x|I)p(\varepsilon|I)$,

$$p(y|x,I)=p(\varepsilon=y-x|I). \tag{4.21}$$

In other words, for this type of additive model, the likelihood is simply given by the probability model of the unpredictable term ε.

In the absence of precise knowledge about the spatio-temporal correlations in the ongoing activity and for inferential purposes only, we assign $\eta(t)$ to be independent identically distributed with a (unknown) time independent diagonal covariance matrix $\mathbf{V}$ with diagonal entries given by σ_η^2, i.e., $\mathbf{V}_{jj}=\sigma^2$. In this way, Equation (4.13) is rewritten as:

$$p(\{s_{nr}(t)\},\{a_{nr}\},\{\tau_{nr}\},\mathbf{B},\sigma_\eta|\{z_{mr}(t)\},I)= \frac{p(\{\eta_{mr}(t)\}|\sigma_\eta,I)\ p(\{s_{nr}(t)\},\{a_{nr}\},\{\tau_{nr}\},\mathbf{B},\sigma_\eta|I)}{p(\{z_{mr}(t)\})}. \tag{4.22}$$

When choosing the model for the likelihood, we apply the principle of maximum entropy. When we have knowledge of the existence of some constraints on the value of a

parameter, we assign a probability that maximizes our uncertainty while accommodating the known constraints. That is, we take into account all the information at our disposal without presupposing information we do not possess. We do this by maximizing the entropy of the probability distribution subject to the known constraints. Thus, given the constraint σ_η^2, we assign to the likelihood the probability density that maximizes the entropy of $\eta(t)$, namely a Gaussian density. The posterior can then be rewritten as

$$p(\{a_{nr}\},\{s_n(t)\},\{\tau_{nr}\},\mathbf{B},\sigma_\eta|\{z_{mr}(t)\},I) \propto p(\{a_{nr}\},\{s_n(t)\},\{\tau_{nr}\},\mathbf{B},\sigma_\eta|I)$$
$$(2\pi\sigma_\eta^2)^{\frac{-MRT}{2}} \exp\left(-\frac{1}{2\sigma_\eta^2}\sum_{m=1}^{M}\sum_{r=1}^{R}\sum_{t=1}^{T}[z_{mr}(t)-\sum_{n=1}^{N}B_{mn}a_{nr}s_n(t-\tau_{nr})]^2\right). \tag{4.23}$$

For simplicity we assume that the matrix **B** and the component parameters a_{nr}, τ_{nr} and $s_n(t)$ are independent and uniformly distributed, with appropriate cutoffs denoting a range of physiologically reasonable values:

$$p(s(t)|I) = const.$$
$$p(a|I) = const., \text{ for } 0 < a \le a_{\max}$$
$$p(\tau|I) = const., \text{ for } \tau_{\min} < \tau \le \tau_{\max}$$
$$p(\mathbf{B}|I) = const.$$

Treating the ongoing noise variance as a nuisance parameter and assigning the Jeffreys prior $p(\sigma_\eta) = \sigma_\eta^{-1}$, we marginalize the posterior over σ_η, i.e.,

$$p(\{a_{nr}\},\{s_n(t)\},\{\tau_{nr}\},\mathbf{B}|\{z_{mr}(t)\},I)$$
$$\int_{-\infty}^{\infty}\sigma_\eta^{-1}(2\pi\sigma_\eta^2)^{\frac{-MRT}{2}}\exp\left(-\frac{1}{2\sigma_\eta^2}\sum_{m=1}^{M}\sum_{r=1}^{R}\sum_{t=1}^{T}[z_{mr}(t)-\sum_{n=1}^{N}B_{mn}a_{nr}s_n(t-\tau_{nr})]^2\right)d\sigma_\eta,$$

(4.24)

obtaining:

$$p(\{a_{nr}\},\{s_n(t)\},\{\tau_{nr}\},\mathbf{B}|\{z_{mr}(t)\},I) \propto p(\{a_{nr}\},\{s_n(t)\},\{\tau_{nr}\},\mathbf{B}|I)$$
$$(2\pi)^{\frac{-MRT}{2}}\Gamma(\frac{MRT}{2})\left(\sum_{m=1}^{M}\sum_{r=1}^{R}\sum_{t=1}^{T}[z_{mr}(t)-\sum_{n=1}^{N}B_{mn}a_{nr}s_n(t-\tau_{nr})]^2\right)^{\frac{-MRT}{2}},$$

(4.25)

where $\Gamma(\frac{MRT}{2})$ is the gamma function evaluated at $\frac{MRT}{2}$. Given the choice of uniform distribution for the components' parameters and dropping the constant term, Equation (4.25) is simply:

$$p(\{a_{nr}\},\{s_n(t)\},\{\tau_{nr}\},\mathbf{B}|\{z_{mr}(t)\},I) \propto \left(\sum_{m=1}^{M}\sum_{r=1}^{R}\sum_{t=1}^{T}[z_{mr}(t)-\sum_{n=1}^{N}B_{mn}a_{nr}s_n(t-\tau_{nr})]^2\right)^{\frac{-MRT}{2}}.$$

(4.26)

Because waveforms, amplitudes and latencies are being estimated simultaneously, there is degeneracy in the model. This problem is eliminated by constraining the norm of each waveform component to equal 1 and the ensemble mean of the latency components to be zero.

Intuition about the characteristics of the MAP solution can be gained by examining the partial derivatives of the posterior with respect to the model parameters. This examination will further lead to a practical and simple estimation algorithm. Let

$$Q=\sum_{m=1}^{M}\sum_{r=1}^{R}\sum_{t=1}^{T}[z_{mr}(t)-\sum_{n=1}^{N}B_{mn}a_{nr}s_n(t-\tau_{nr})]^2. \qquad (4.27)$$

Then the logarithm of the posterior can be written as:

$$\ln P = -\frac{MRT}{2}\ln Q + const. \tag{4.28}$$

For the partial derivatives we employ the specific indices i, j, p, q corresponding to generic indices m, n, r, and t, respectively. The first partial derivative with respect to $s_j(q)$ is:

$$\frac{\partial \ln P}{\partial s_j(q)} = -\frac{MRT}{2} Q^{-1} \frac{\partial Q}{\partial s_j(q)}, \tag{4.29}$$

where

$$\frac{\partial Q}{\partial s_j(q)} = -2\sum_{m=1}^{M}\sum_{r=1}^{R}[WB_{mj}a_{jr} - (B_{mj}a_{jr})^2 s_j(q)], \tag{4.30}$$

and

$$W = z_{mr}(q+\tau_{jr}) - \sum_{\substack{n=1 \\ n\neq j}}^{N} B_{mn}a_{nr}s_n(q-\tau_{nr}+\tau_{jr}). \tag{4.31}$$

Solving $\frac{\partial Q}{\partial s_j(q)} = 0$ for the optimal parameters $\hat{s}_j(q)$ results in:

$$\hat{s}_j(q) = \frac{\sum_{m=1}^{M}\sum_{r=1}^{R} WB_{mj}a_{jr}}{\sum_{m=1}^{M}\sum_{r=1}^{R}(B_{mj}a_{jr})^2}. \tag{4.32}$$

The above equation does not have a closed form solution since the right hand side depends on the other estimated parameters. However, intuition about the type of the solution can be obtained by examining the term W. Basically, it involves: the time shifting of the data according to the latencies of the estimated component, i.e. $z_{mr}(q+\tau_{jr})$; the

subtraction from the data of the other scaled and time shifted components, i.e. $B_{mn}a_{nr}s_n(q-\tau_{nr}+\tau_{jr})$, $n \neq j$; and the average of the scaled residuals, where the scaling is given by the terms $B_{mj}a_{jr}$. Similarly, for the parameters a_{jp} and B_{ij} we obtain:

$$\hat{a}_{jp} = \frac{\sum_{m=1}^{M}\sum_{t=1}^{T} UV}{\sum_{m=1}^{M}\sum_{t=1}^{T} V^2}, \tag{4.33}$$

where $U = z_{mp}(t) - \sum_{\substack{n=1 \\ n\neq j}}^{N} B_{mn}a_{nr}s_n(t-\tau_{np})$ and $V = B_{mj}s_j(t-\tau_{jp})$; and

$$\hat{B}_{ij} = \frac{\sum_{r=1}^{R}\sum_{t=1}^{T} ZY}{\sum_{r=1}^{R}\sum_{t=1}^{T} Y^2}, \tag{4.34}$$

where $Z = z_{ir}(t) - \sum_{\substack{n=1 \\ n\neq j}}^{N} B_{in}a_{nr}s_n(t-\tau_{nr})$ and $Y = a_{jr}s_j(t-\tau_{jr})$.

Notice that the formulas derived for $\hat{a}_{jp}$ and $\hat{B}_{jp}$ are related to matching filter solutions. For example, $\hat{a}_{jp}$ is given by the projection of the detector-scaled component, i.e. V, and the data after subtraction of the other scaled and time shifted components, i.e. U. Also, the estimation of a_{jp} does not directly depend on the trials $r \neq p$.

For the latency parameter, setting $\frac{\partial Q}{\partial \tau_{jp}} = 0$ leads to the following equation[3]:

---

[3] To deal with continuous latency values, we need a continuous time model of the component waveform, which can be implemented using a spline model, for example.

$$2\sum_{m=1}^{M}\sum_{t=1}^{T}\left([z_{mp}(t)-\sum_{\substack{n=1\\n\neq j}}^{N}B_{mn}a_{np}s_n(t-\tau_{np})]B_{mj}a_{jp}s_j'(t-\tau_{jp})-a_{jp}^2s_j'(t-\tau_{jp})s_j(t-\tau_{jp})\right)=0,$$

(4.35)

where $s_j'(t-\tau_{jp})$ is the time derivative of $s_j'(t)$ evaluated at $(t-\tau_{jp})$. The solution for $\hat{\tau}_{jp}$ is more difficult since the latencies appear as an implicit parameter in the waveform function. An alternative and more intuitive solution can be obtained by directly examining the condition for maximization of the logarithm of the posterior, which is equivalent to minimization of the quadratic term Q, Equation (4.27). Expansion of this term results in:

$$Q=\sum_{n=1}^{M}\sum_{r=1}^{R}\sum_{t=1}^{T}\left(z_{mr}^2(t)+[\sum_{n=1}^{N}B_{mn}a_{nr}s_n(t-\tau_{nr})]^2-2z_{mr}(t)\sum_{n=1}^{N}B_{mn}a_{nr}s_n(t-\tau_{nr})\right). \quad (4.36)$$

As τ_{jp} is varied, only the cross-terms in $z_{mp}(t)\sum_{n=1}^{N}B_{mn}a_{np}s_n(t-\tau_{np})$ corresponding to j^{th} component change (as long as the evoked components $s_n(t)$ are zero outside a closed time interval $[t_0,t_f]$). Notice that the estimation of τ_{jp} does not directly depend on the trials $r\neq p$. Thus the optimal parameter $\hat{\tau}_{jp}$ is found by maximizing

$$\rho(\tau)=\sum_{m=1}^{M}\sum_{t=1}^{T}\left(B_{mj}a_{jp}s_j(t-\tau)[z_{mp}(t)-\sum_{\substack{n=1\\n\neq j}}^{N}B_{mn}a_{np}s_n(t-\tau)]\right), \quad (4.37)$$

which if properly normalized is just the cross-correlation between the estimated component and the data after the contributions from the other components has been subtracted off. This is then averaged over all the detectors. Thus:

$$\hat{\tau}_{jp} = \arg\max \rho(\tau). \tag{4.38}$$

The multi-channel generative model for local field potentials is particularly suited for closely positioned electrodes. An example where that is the case is the cortical intra-laminar recordings (see Figure 4.1). Work with this type of data set is currently in development (Shah, et al., 2001). For the specific data set considered in this dissertation, the mixture of distant sources of LFP activity into the same channel recording is unlikely. For this reason in the algorithm implementation to be presented next, we undertake the estimation of single trial parameters on single channel data sets.

4.2 Algorithm Implementation

We have experimented with several approaches for the exploration of the posterior probability. The MCMC is the most informative of all, but it has the major disadvantage of being computationally intensive, together with the fact that it is not easy in practice to establish rigorously when the system has reached equilibrium and when to stop the simulation. When searching exclusively for the MAP solution, we also employed conjugate gradients and quasi-Newton trust region methods, since the partial derivatives and the Hessian matrix of the posterior probability are easily derived. The optimization does not need to be done for all the parameters at once, but in blocs, e.g., first optimize with respect to the latencies, then the waveforms, and so on. These are less computationally intensive, but still require multiple searches starting at different initial conditions or simulated annealing to avoid local maxima. In the simplest case of applying the Newton-Raphson to the problem of finding the MAP solution of the parameter θ for

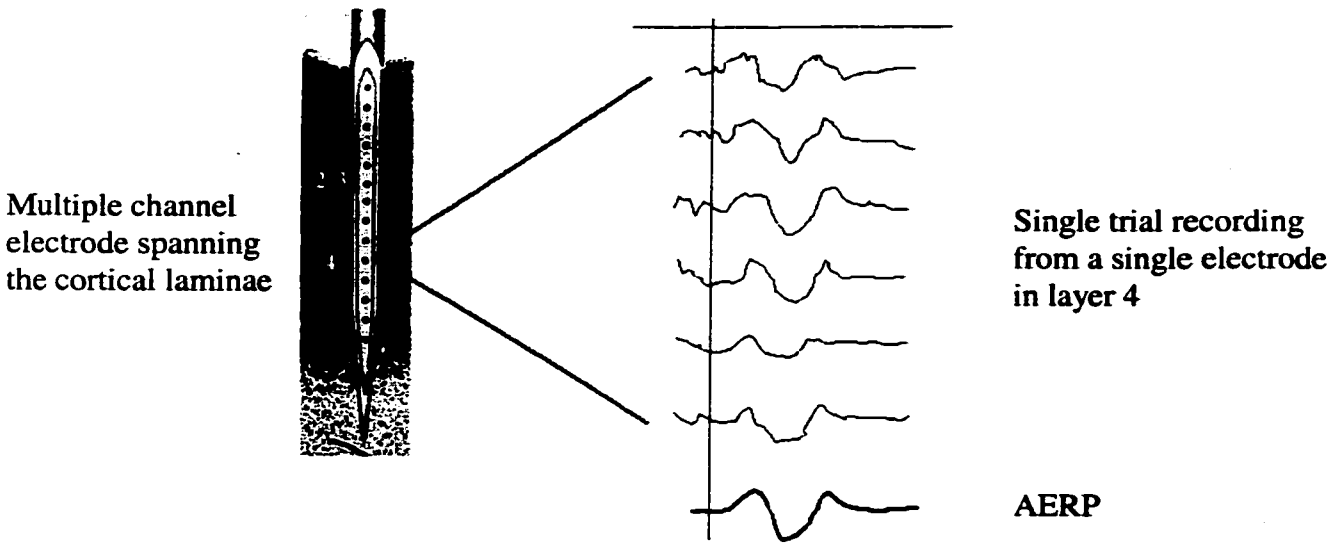

Figure 4-1: Potential application of the multi-channel multicomponent event-related LFP estimation. Local field potentials recorded from very close sites will likely lead to mixture of sources in the recordings of a single channel. Similar potential application would be to recordings from horizontal intracortical electrode arrays.

the logarithm of the posterior probability $L(\theta)$, the Newton-Raphson solution after the m^{th} iteration is given by

$$\theta^{m+1} = \theta^m - [\nabla\nabla L(\theta^m)]^{-1} \nabla L(\theta^m), \tag{4.39}$$

where $\nabla\nabla L(\theta^m)$ is the Hessian matrix whose entries are given by

$$\nabla\nabla L(\theta^m)_{ij} = \left.\frac{\partial^2 L(\theta)}{\partial\theta_i \partial\theta_j}\right|_{\theta^m}, \tag{4.40}$$

and $\nabla L(\theta^m)$ is a vector whose components are specified by

$$\nabla L(\theta^m)_i = \left.\frac{\partial L(\theta)}{\partial\theta_i}\right|_{\theta^m}. \tag{4.41}$$

As a side note, $-[\nabla\nabla L(\theta_0)]^{-1}$ is a quadratic approximation of the local curvature around the mode θ_0 and, for roughly symmetric pdfs, can be used as an estimated of the covariance matrix. For Gaussian pdfs, this result is exact. Also, if $L(\theta)$ is the logarithm of the likelihood and $\hat{\theta}$ is the MAP estimate, $-[\nabla\nabla L(\hat{\theta})]^{-1}$ is called the observed Fisher information matrix.

Better stability properties are achieved by introducing a small negative constant c:

$$\theta_{P+1} = \theta_P - [\nabla\nabla L(\theta_P) + c\mathbf{I}]^{-1} \nabla L(\theta_P). \tag{4.42}$$

where $\mathbf{I}$ is the identity matrix. The result is better than multiplying the matrix-vector product on the right of Equation (4.39) by a fractional constant. The ultimate effect of c is, while preserving the same eigenvectors of $\nabla\nabla L(\theta^m)$, to change the eigenvalues in such a way that it decreases the contributions of the eigendirections associated with large uncertainty (Sivia, 1996).

However, given the analysis in the previous section, an even simpler heuristic algorithmic implementation can be designed when dealing with high signal-to-noise ratio recordings. It consists also of an iterative algorithm where, after starting with an initial guess for all the parameters, they are updated in blocs at each iteration step. The latencies τ_{nr} are implemented in terms of integer time offsets. More specifically, for a single channel data set $\{z_r(t)\}$:

(0) At the iteration $m = 1$, the initial guess for the amplitudes and time offsets are set to $a_{nr}^1 = 1, \tau_{nr}^1 = 0, \ \forall n, r$. The number of components N and an initial guess for their waveforms $s_n^1(t)$ can be specified based on the inspection of the AERP (see Figure 4-2).

The iteration cycle is given by the following 4 steps:

(1) For all the trials, estimate the single trial time offsets for the first up to the N^{th} component according to $\hat{\tau}_{jp}^{m+1} = \arg\max \rho^m(\tau)$. For the single channel case, this corresponds to maximizing

$$\rho^m(\tau) = \sum_{t=1}^{T} \left(a_{jp}^m s_j^m(t-\tau)[z_p(t) - \sum_{\substack{n=1 \\ n \neq j}}^{N} a_{np}^m s_n^m(t-\tau)] \right).$$

The cross-correlation can be employed, and if the components are not highly overlapped, it can be computed directly between a single component waveform and the raw single trial data.

More specifically, let $k_a \le k \le k_b$ be a post-stimulus time segment where $s_n^m(k)$ is nonzero, and k is an integer index for the sampled data point. Here k_a and k_b are indices for the first and last data point of the segment. For convenience, we write the latency term τ_{np} as an argument for $z_p(t)$. The cross-correlation between the specified segment of the estimated evoked response and a corresponding segment of single trial data, i.e. $\frac{< s_n^m(k) z_p(k+\tau_{np}^m) >_k}{\sqrt{< [s_n^m(k) - < s_n^m(k) >_k]^2 >_k < [z_p(k+\tau_{np}^m) - < z_p(k+\tau_{np}^m) >_k]^2 >_k}}$, is then computed for different integer lags τ, with $k_0 + \tau_{np} > 0$ for k_0 corresponding to the sample index at stimulus onset. Preferentially, the data should have been sampled or resampled at 1KHz or higher. The lag that gives the maximum positive cross-correlation value is then chosen to be the latency estimate $\hat{\tau}_{jp}^{m+1}$. The $\hat{\tau}_{jp}$ obtained in this way represents the number of data point shifts with respect to the position in time of the component waveform $s_n(k)$. Given an approximate knowledge of where in time the component is expected to happen, an interval for the search of the optimal latency $\hat{\tau}_{jp}$ can be applied. In this way, the possibility that the component matches by chance the waveform of unrelated ongoing activity is diminished.

(2) Estimate the waverforms according to:

$$\hat{s}_j^{m+1}(q) = \frac{\sum_{r=1}^{R} W^m a_{jr}^m}{\sum_{r=1}^{R} (a_{jr}^m)^2},$$

with $W^m = z_r(q+\tau_{jr}^m) - \sum_{\substack{n=1 \\ n\neq j}}^{N} a_{nr}^m s_n^m(q-\tau_{nr}^m+\tau_{jr}^m)$.

(3) For all the trials and components, estimate the amplitudes according to:

$$\hat{a}_{jp}^{m+1} = \frac{\sum_{t=1}^{T} U^m V^m}{\sum_{t=1}^{T} [V^m]^2},$$

with $U^m = z_p(t) - \sum_{\substack{n=1 \\ n\neq j}}^{N} a_{np}^m s_n^m(t-\tau_{np}^m)$ and $V = s_j^m(t-\tau_{jp}^m)$. Notice that since the norm of each component waveform is being constrained to equal 1, $\sum_{t=1}^{T} [V^m]^2 = 1$.

(4) Repeat (1) through (3) for M iterations.

A few iterations are usually employed to avoid stability problems. In the data set employed here, where the overlap between the components is not large and signal-to-noise ratio is high, two iterations already give reasonable results.

4.3 Separation of Event-Related Components and Ongoing Activity: Identification of the Origin of Event-Related Increase in Ensemble Variance Time Function.

As verified in Chapter 3, a common phenomenon observed in the data set is the existence of event-related increase in the ensemble variance time function, defined as:

$$\sigma_z^2(t) = < [z_r(t) - < z_r(t) >_r]^2 >_r. \qquad (4.43)$$

For illustration, we choose a parietal channel recording time series from the monkey GE, with GO response to one single stimulus pattern. Inspection of the AERP suggests the

presence of two components (see Figures 4-2). The AERP also provides the initial guess for the components' waveforms, as well as initial information about the range for the single trial latency values. The resulting estimated waveforms for the two components, and the histogram for their respective single trial amplitude and latencies after two iterations of the estimation algorithm are shown in Figure 4-3. Figure 4-4 provides a few examples of the single trial reconstructed event-related phase locked components superposed onto the single trial data.

Once the single trial event-related components have been estimated, the ongoing activity and its contribution to the ensemble variance are obtained according to:

$$\hat{\eta}_r(t) = z_r(t) - \sum_{n=1}^{N} \hat{a}_{nr}\hat{s}_n(t - \hat{\tau}_{nr}), \tag{4.44}$$

$$\sigma^2_{\hat{\eta}}(t) = < [\hat{\eta}_r(t) - < \hat{\eta}_r(t) >_r]^2 >_r. \tag{4.45}$$

with $N = 2$. Figure 4-5 shows the estimated ensemble mean and variance time function for the estimated ongoing activity.

The differential variability in amplitude and latency of the event-related components can be related to behavioral variables. An example in Figure 4-6 is provided for the same parietal channel recording showing that latter components present higher correlation with the reaction time (RT). Another example of removal or the transient increase in ensemble variance time function for a striate site is provided in Figure 4-7.

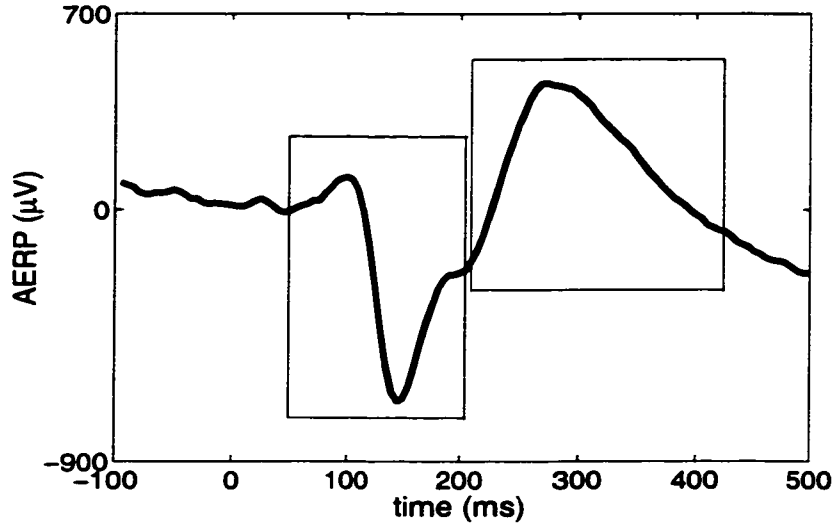

Figure 4-2: Estimation of single trial event-related components. The AERP from the recordings of a parietal channel suggests the modeling of two main event-related components initially represented by the waveforms enclosed by the rectangles. The first and second components happen around 100 and 250 ms, respectively. The initial guesses for the components waveforms are obtained by the AERP waveforms in the selected regions (the waveforms are set to zero for times outside the rectangular selection in the first iteration). The AERP also suggests a range of latency values for each of the components. Ensemble data: 222 trials.

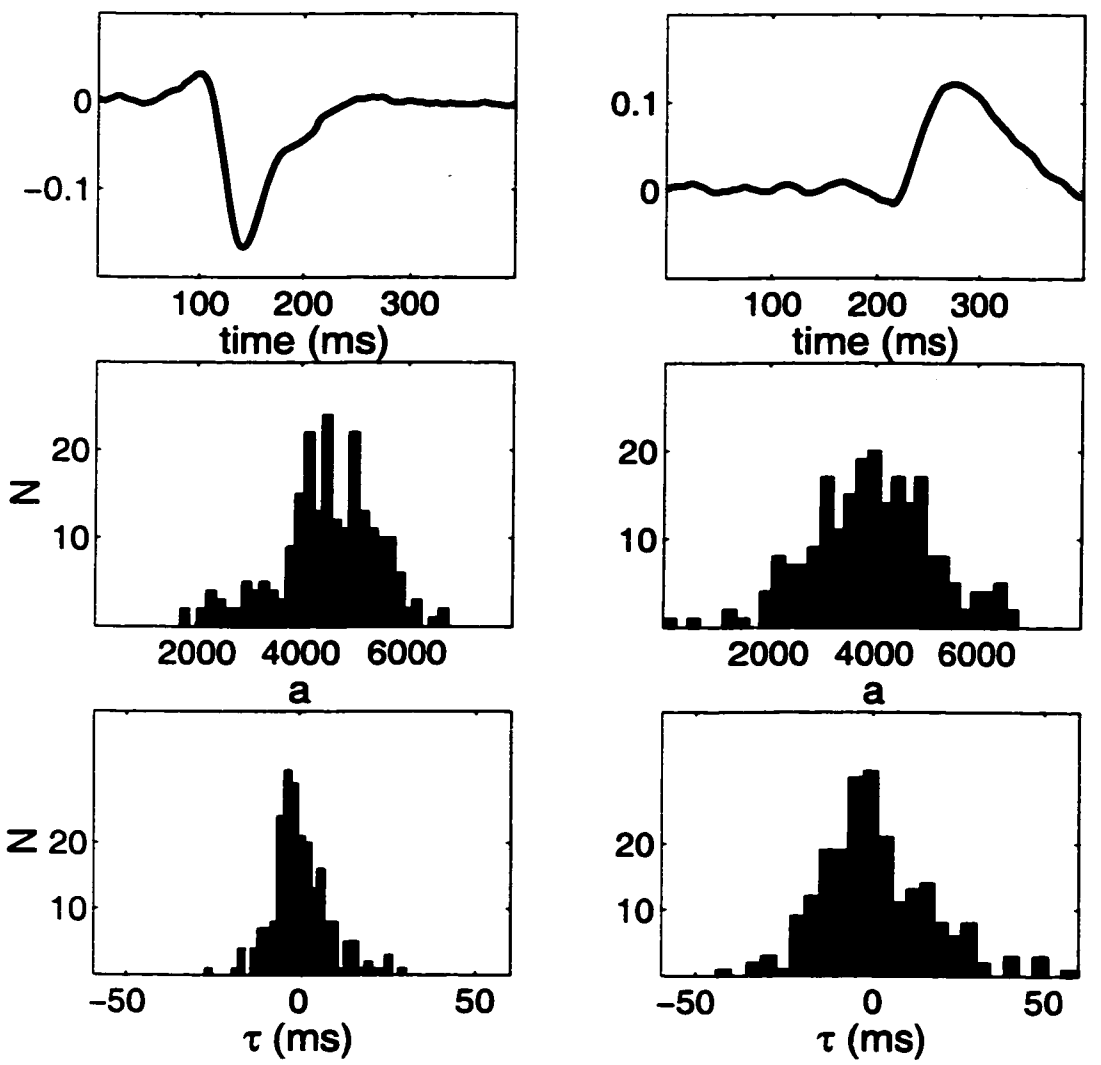

Figure 4-3: Estimated single trial event-related components. The estimated component waveforms for the same parietal channel recording as in Figure 4-2 are shown in the top row. The waveforms have been normalized by constraining their norm to equal 1. In the middle and bottom row, histograms for the single trial amplitudes and latencies for the respective components are shown. Differential variability for amplitude and latency between the two components is observed: notice that the later component has larger amplitude and latency variability. Single trial parameters were obtained after two iterations. Latency values are relative to the largest extrema of the corresponding estimated component waveform.

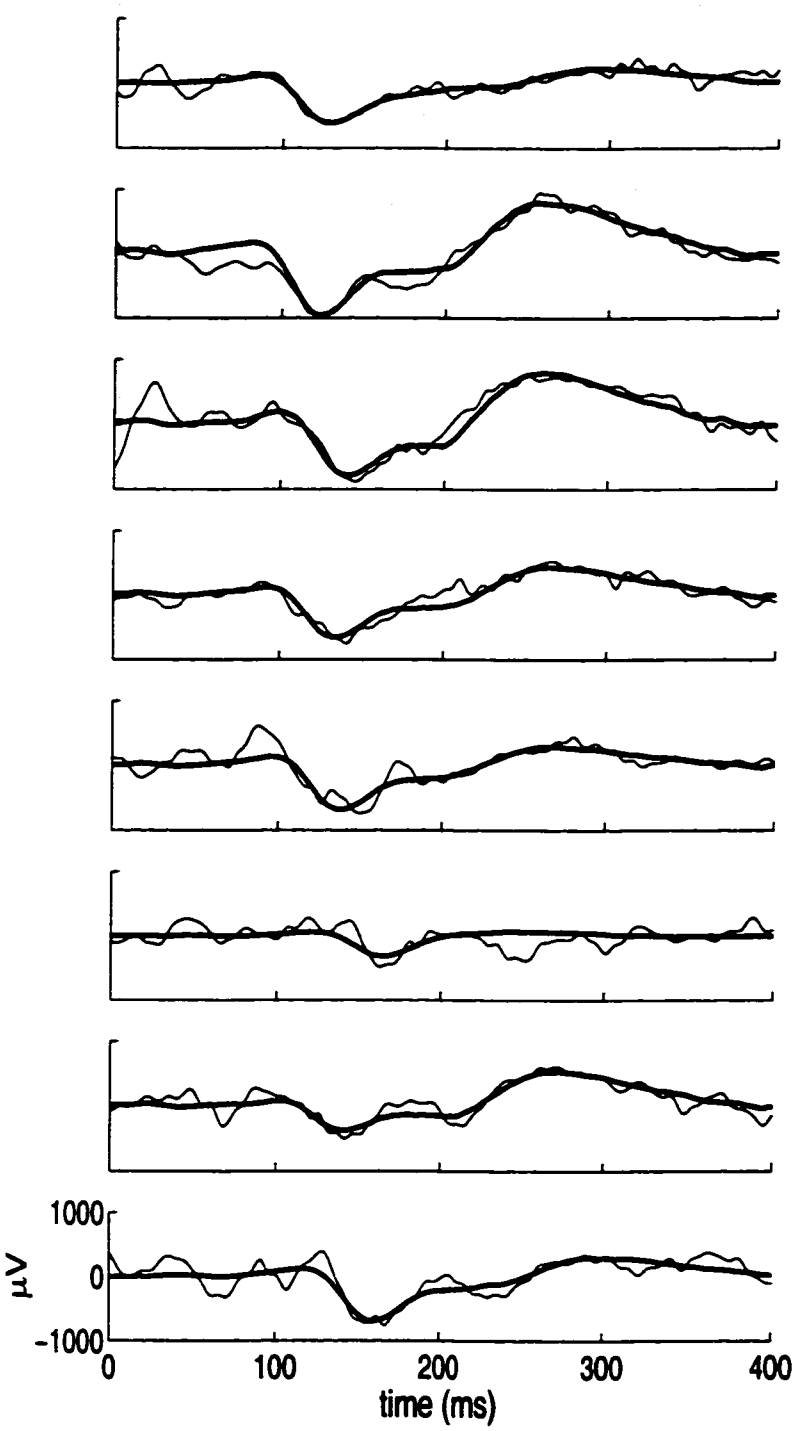

Figure 4-4: Single trial estimated event-related phase-locked components. Single trial examples of the two estimated components for the same parietal channel (Figure 4-2) are shown by the thick curve superposed on the respective single trial data.

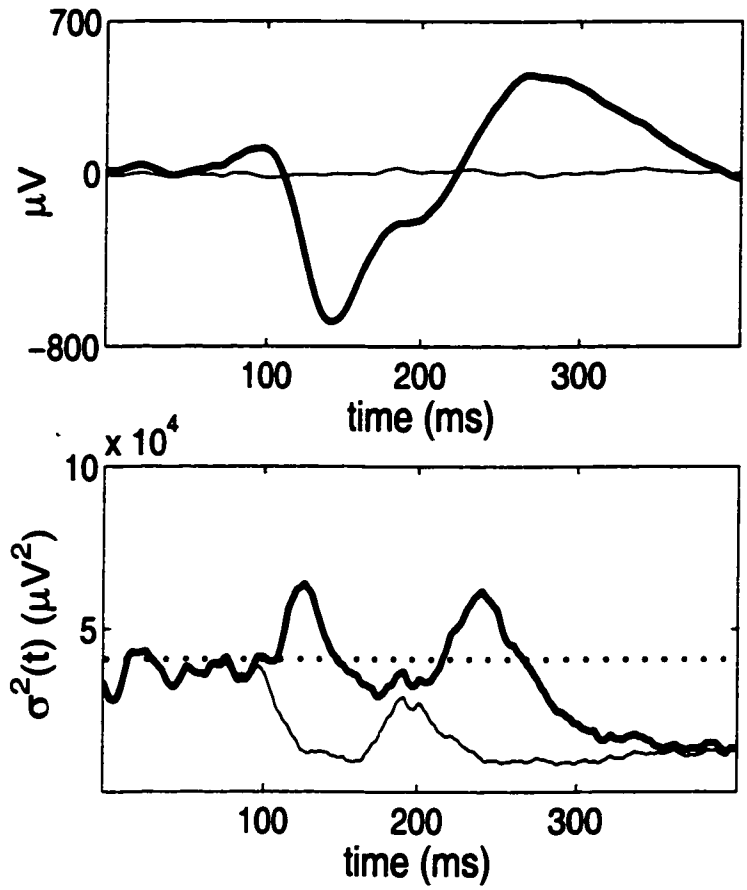

Figure 4-5: Estimation of ongoing activity. Top plot: the AERP or ensemble average of the original time series for a parietal channel (monkey GE) is shown by the thick curve. The single trial ongoing activity time series is obtained by subtracting the estimated single trial event-related components from the corresponding original single trial time series. The ensemble mean of the estimated ongoing time series is shown by the thin curve. As expected for a good estimation (see the amplitude and latency distributions in Figure 4-3), it is approximately zero. Bottom plot: the ensemble variance time function of the original time series is given by the thick curve. During the event-related response period, there is a clear modulation of the variance compared to variance level before 100 ms (pre-stimulus or pre-sensory activation period) which is seen as two transient peaks. The doted line corresponds to the mean variance level in the pre-stimulus period. The ensemble variance time function of the estimated ongoing activity is shown by the thin curve. Two main aspects are noticed. First, the transient increases in variance have been removed. Second, the variance of the ongoing activity is also clearly nonstationary, decreasing shortly after the initial sensory activation (~ 100 ms); the variance of the ongoing time series fluctuates around a baseline that is about the same level of the variance of the original time series after the event-related components are almost over, i.e., after ~ 300 ms. This second aspect and the approximately zero mean of the estimated

ongoing activity indirectly imply a good estimation of the single trial event-related components. Stimulus onset is at 0 ms.

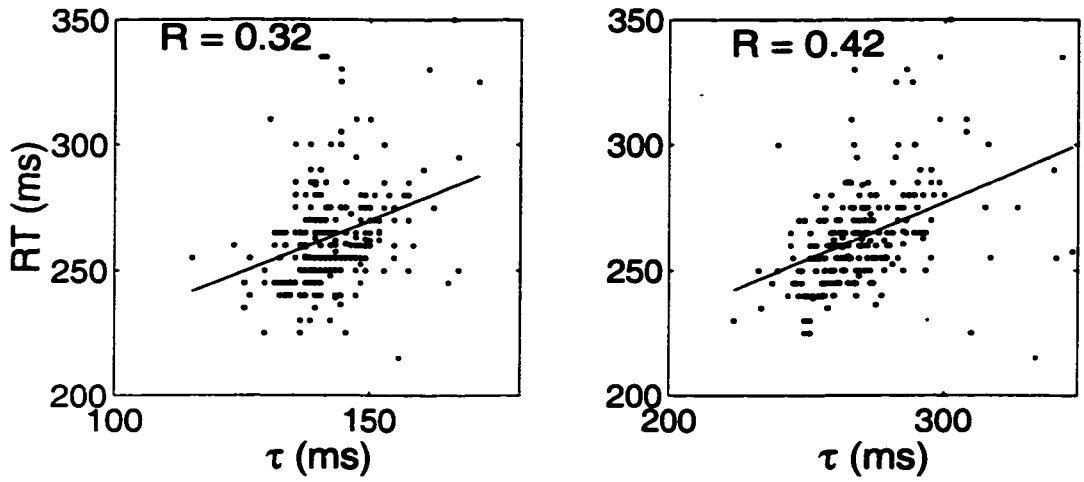

Figure 4-6: Relations between components' latency and reaction time. The component latency and the reaction time RT seem to become progressively more correlated as the component happens closer to the motor response.

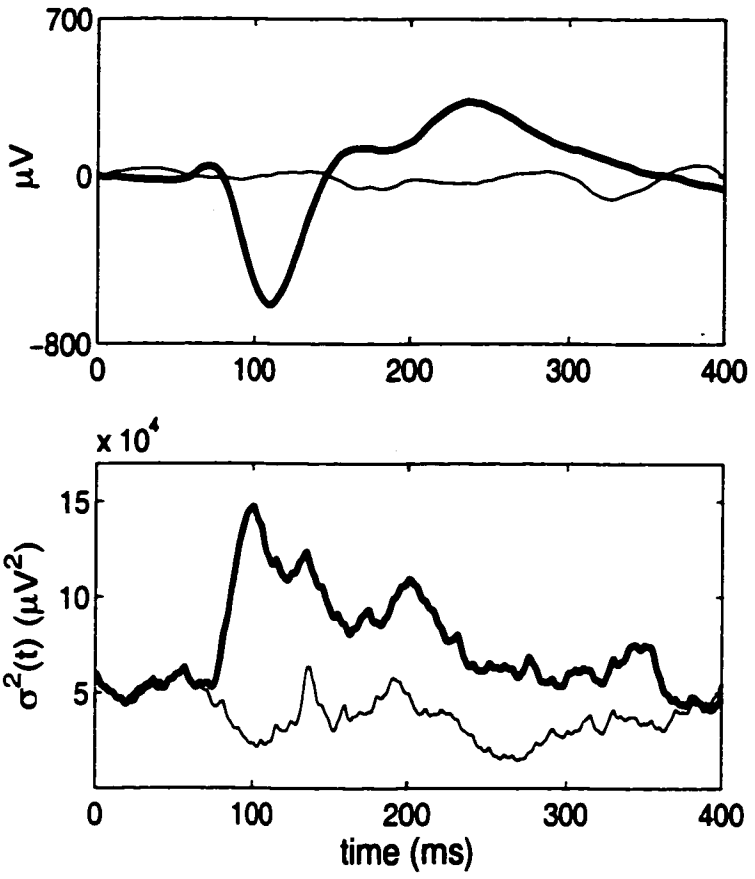

Figure 4-7: Estimation of ongoing activity. Top plot: the AERP or ensemble average of the original time series for a striate channel recordings (monkey GE) is shown by the thick curve. The single trial ongoing activity time series is obtained by subtracting the estimated single trial event-related components from the corresponding original single trial time series. The ensemble mean of the estimated ongoing time series is shown by the thin curve. Bottom plot: the ensemble variance time function of the original time series is given by the thick curve. The estimation of the ongoing activity was based on the estimation of two event phase-locked components, one around 110 ms and the other around 240 ms. Stimulus onset is at 0 ms.

The estimation of single trial amplitudes and latencies gives the opportunity to search for the origins of this trial-to-trial variability. One possibility is that it arises from multiplicative effects between ongoing and evoked activities. To investigate this alternative, the correlation between the single trial pre-stimulus power spectrum at different frequency bands and the single trial amplitudes for the N1 components of the same previously employed parietal and striate channels (monkey GE) were calculated. The power spectrum for each single trial pre-stimulus time series (from –100 to 50 ms, stimulus onset at time zero) was computed and integrated over 4 frequency bands: alpha (7 - 14 Hz), beta 1 (14 – 20 Hz), beta 2 (20 - 30 Hz) and gamma (30 – 80 Hz). The power spectrum was obtained after the FFT of the Hanning windowed data segments. For the computation of the correlation coefficients, the DC component was also included. The results are shown in Table 2, and point to a very small if any linear relationship between the pre-stimulus power and the single trial parameters. The strongest correlations were obtained for the parietal channel: a positive correlation between the DC component and both the amplitude and the latency of the N1 component.

Table 4.1: Correlation between single trial parameters and single trial pre-stimulus power

		DC	Alpha	Beta 1	Beta 2	Gamma
Striate	A	-0.10	-0.13*	-0.10	-0.07	-0.13
	τ	-0.06	0.05	-0.01	-0.02	-0.04
Parietal	A	0.33*	0.07	-0.13*	-0.08	-0.11
	τ	0.21*	0.05	0.04	0.00	-0.09

*: significant correlation coefficient, $p<0.05$.

4.4 Removing the Effect of Trial-to-Trial Variability of Event Phase-Locked Responses from the Statistical Measures of Interdependence.

Theoretical considerations and experimental results given in Chapters 1, 2 and 3 have shown that characteristic fluctuations in the variance of remnant event-related phase locked components in the residual time series underlie temporal modulation of interdependence measures, the referred signal-to-noise effect. Thus, once the single trial event-related components are well estimated and removed from the time series data set, the statistical measures should not be affected by the signal-to-noise ratio effect anymore. Furthermore, if the previously observed event-related modulation of coherence time functions is significantly attenuated when computed on this new residual time series, it would imply that most of the observed modulation was indeed originated from the signal-to-noise ratio fluctuations.

To test this possibility, we estimated the single trial event-related phase-locked components for every channel recordings from the four monkeys in the same data used in Chapter 3. Since the experimental results were focused on the event-related modulation of 12 Hz coherence happening around 120 ms, the estimation was restricted to the single trial component occurring about this time. In other words, the number of components was set to 1. The spectral coherence time functions were computed as before on the new residual time series defined as:

$$\hat{\xi}_{mr}(t) = \hat{\eta}_{mr}(t) - < \hat{\eta}_{mr}(t) >_r. \tag{4.46}$$

The comparison between the 12Hz coherence modulation in the original residual time series and the 12 Hz coherence computed on the new residual time series is made through

the use of histograms. Histograms of maximum 12 Hz coherence values during the post-stimulus period (0-200 ms) were computed for all channel pairs for each subject, both on the original residual time series ensembles and on the new residual ensembles. The results are shown in Figure 4-8. In summary, with the original residual time series, 88% (GE), 71% (LU), 65% (TI) and 53% (PE) of the post-stimulus (0-200 ms) coherence maxima for each channel pair are greater than 0.1, compared to only 9% (GE), 5% (LU), 6% (TI) and 18% (PE) for the coherence maxima computed on the new residual time series ensembles.

4.5 Discrimination Between Simultaneous Cortical Processes

As a final application, the single trial component estimation is used to investigate different cortical processes recorded by different channels, which from the inspection of the respective AERP seem to happen simultaneously, but that given the information provided by the AERPs only, it is not certain if and how the processes relate to each other or to behavioral variables. An exemplar case is chosen from a striate and a somatosensory channel from the monkey GE (response GO, stimulus pattern DIAMOND). The results are given in Figure 4-9, where it is clear that although the AERPs show component activations in the striate and somatosensory happening almost simultaneously just before the mean reaction time, the one in the somatosensory is correlated with the single trial reaction time, while the one in the striate is not. This indicates that the activations are related to different and perhaps largely independent processes.

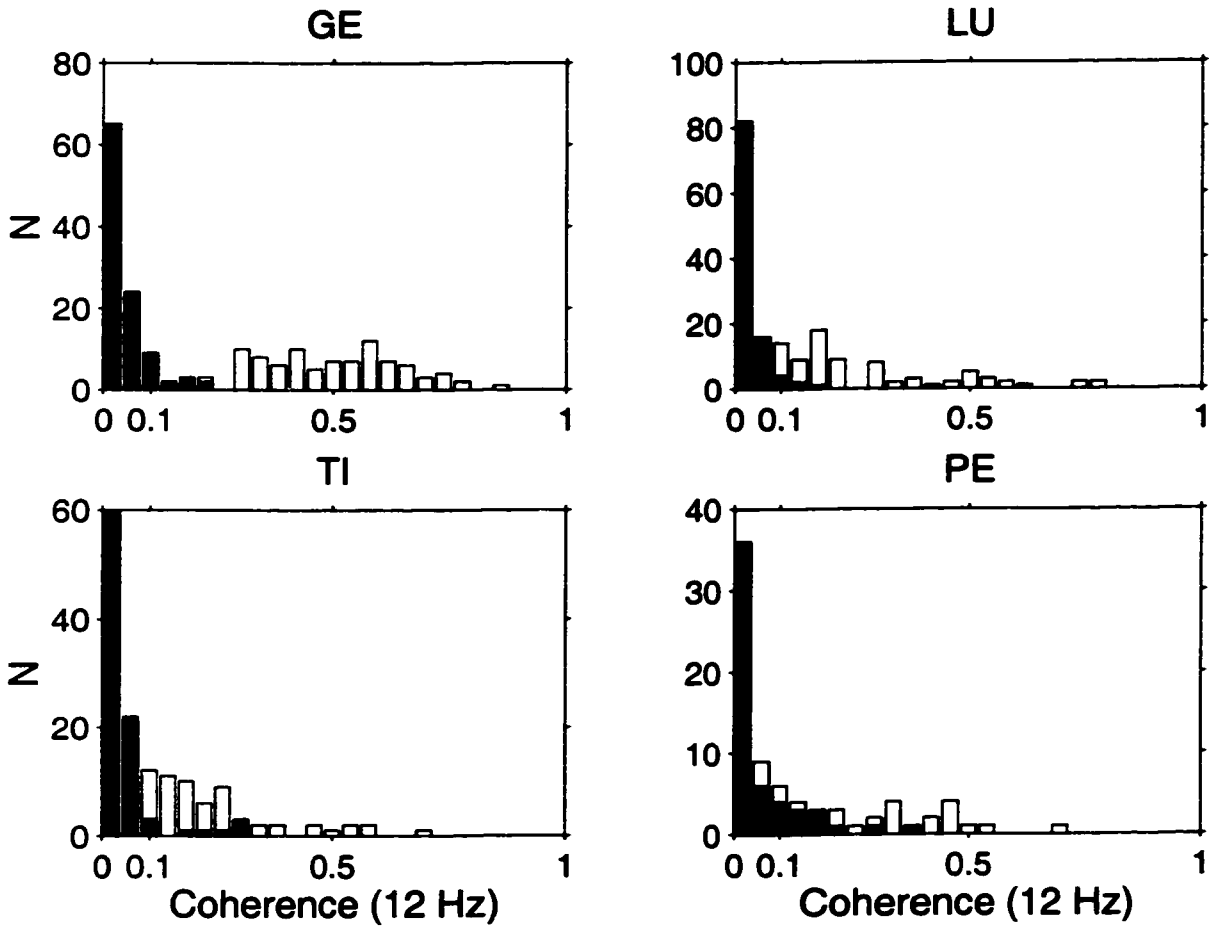

Figure 4-8: Removal of the trial-to-trial variability effect on the coherence time functions. Each plot shows the histogram for the maximum 12Hz coherence value during the post-stimulus period (0-200 ms) for all the channel pair combinations from a single subject. Histograms were computed on the original residual time series ensemble (open bars) and on the new residuals, i.e., the single trial event-related-response subtracted time series ensemble (solid bars). As can be seen, after the single trial responses have been subtracted from the time series, the majority of the coherence maximum values are close to zero. More specifically, for the original residual time series: 88% (GE), 71% (LU), 65% (TI) and 53% (PE) of the post-stimulus coherence maxima for each channel pair are greater than 0.1; while this percentage decreases to: 9% (GE), 5% (LU), 6% (TI) and 18% (PE) when coherences are computed on the new residuals. Squared coherence values were employed.

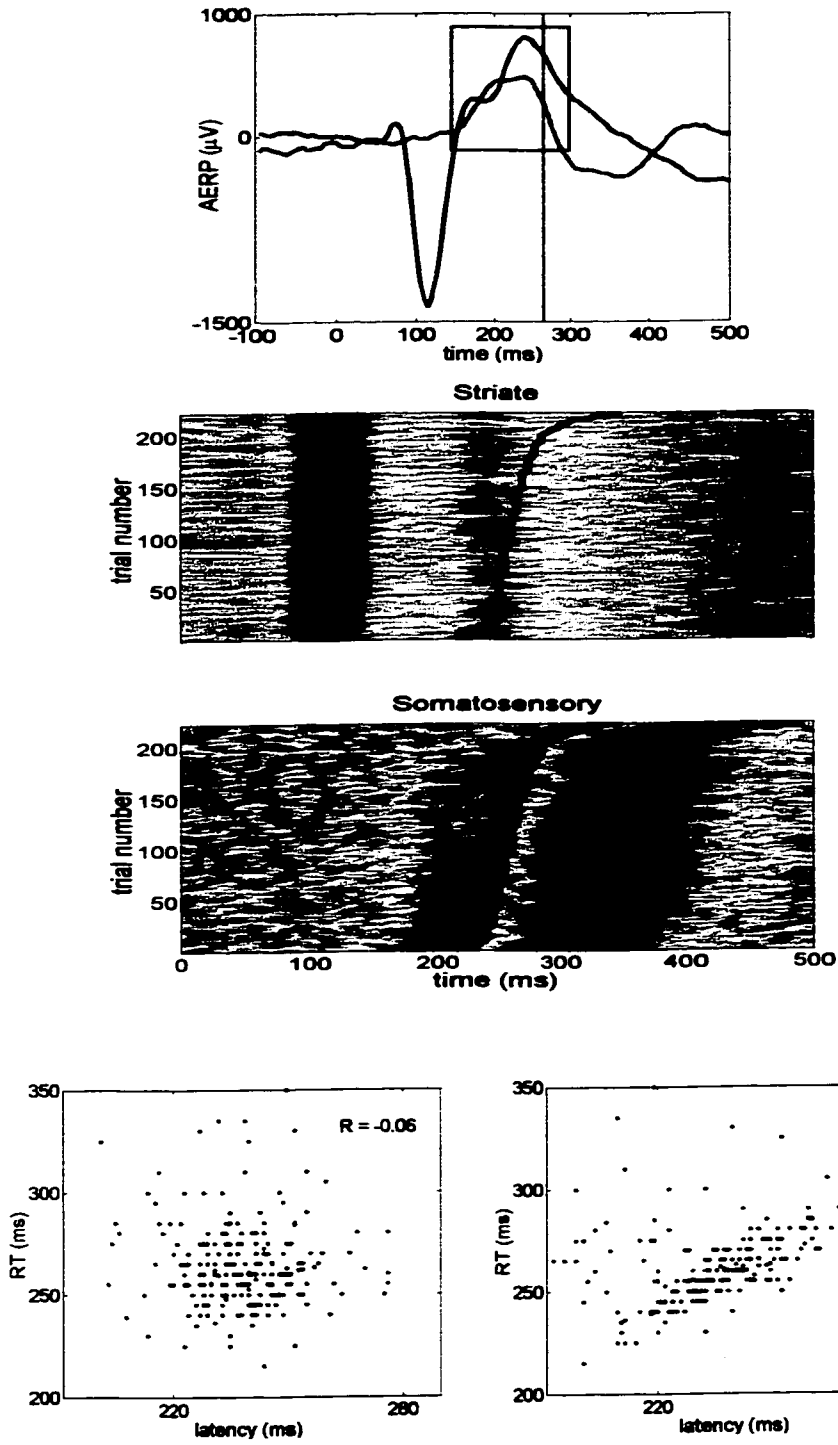

Figure 4-9: Discrimination between processes related to two simultaneous events. Top plot: the AERPs of a striate (str3) and a somatosensory (som) channel (red curve) are shown. The two AERPs contain an event-related component happening just before the mean reaction time (vertical line at time = 265 ms). The components are indicated by the

rectangles. Middle plot: the normalized single trial time series from the two channels are sorted according to the reaction time of the GO response in the corresponding trial. The gray scale denotes the amplitude of the LFPs. The RT values in each trial are shown by the superposed black curve. Inspection of these plots suggests a strong correlation between the event-related phase-locked component for the somatosensory channel, but not for the striate channel. Bottom plots: the scatter plots shows the estimated single trial latencies versus the RT for the striate channel (left) and somatosensory channel (right), as well as the correlation coefficients for the latencies and RT. This more detailed analysis clearly shows that the single trial latencies for the striate component are not linearly correlated with the RT. This finding indicates that the two components observed in the AERPs of the striate and somatosensory channels relate to different cortical processes, even though they happen about the same time in the behavioral task.

4.6 Generative Model in the Frequency Domain

The inferential framework presented here has treated the ongoing activity as a white process. Introducing knowledge of the temporal correlations of the ongoing activity in the time domain model can be very cumbersome, since the correlations of many different time lags and order need to be carried in the derivations. An alternative formulation of the generative model in the frequency domain can then be useful when dealing with 2nd order temporal correlations. The main advantage of the frequency domain formulation is that the correlation functions are diagonal in the frequency space for stationary processes. This leads to simplifications in the derivations.

Expressing the model in Equation (4.12) in the frequency domain, we obtain

$$z_{mr}(\omega) = \sum_{n=1}^{N} B_{mn} a_{nr} s_n(\omega) e^{-i\omega\tau_{nr}} + \eta_{mr}(\omega), \qquad (4.47)$$

where $\omega = [0, 2\pi/T, 4\pi/T, \ldots]$, and $z_{mr}(\omega)$, $s_{nr}(\omega)$ and $\eta_{mr}(\omega)$ are the Fourier transforms of the data, event-related phase locked signals and ongoing activity, respectively.

For simplicity, we assume that the ongoing processes are independent across channels. In this fashion we only need to carry the power spectrum density $\Gamma_m(\omega)$, instead of the whole spectral matrix $S(\omega)$, in the formulation. To begin with, consider the case where $\Gamma_m(\omega)$ is known, as for example when the ongoing activity is WWS and knowledge of its second order statistics is satisfactorily estimated, for example, from the pre-event recordings. The posterior probability is then expressed as:

$$p(\{s_n(\omega)\},\{a_{nr}\},\{\tau_{nr}\},\mathbf{B}|\{z_{mr}(\omega)\},\{\Gamma_m(\omega)\},I) =$$

$$\frac{p(\{z_{mr}(\omega)\}|\{s_n(\omega)\},\{a_{nr}\},\{\tau_{nr}\},\mathbf{B},\{\Gamma_m(\omega)\},I)\; p(\{s_n(\omega)\},\{a_{nr}\},\{\tau_{nr}\},\mathbf{B}|I)}{p(\{z_{mr}(\omega)\})} \tag{4.48}$$

As before a Gaussian likelihood is assigned, i.e., the ongoing activity $\eta_m(\omega)$ is treated as independent (across frequency) complex, normal variable with zero mean and variance given by $\Gamma_m(\omega)$ (Brillinger, 1975). The posterior probability is:

$$p(\{s_n(\omega)\},\{a_{nr}\},\{\tau_{nr}\},\mathbf{B}|\{z_{mr}(\omega)\},\{\Gamma_m(\omega)\},I) \propto p(\{s_n(\omega)\},\{a_{nr}\},\{\tau_{nr}\},\mathbf{B}|I)$$

$$\prod_{mr\omega}[2\pi\Gamma_m(\omega)]^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\Gamma_m(\omega)^{-1}\left|z_{mr}(\omega) - \sum_n^N B_{mn}a_{nr}s_n(\omega)e^{-i\omega\tau_{nr}}\right|^2\right]$$

(4.49)

Notice that if $\Gamma_m(\omega) = const.$ for all ω, the derivations should be equivalent to the one in time domain. The MAP solution should not change by introducing knowledge of the temporal structure of the ongoing activity, since the power spectrum (or even the whole spectral matrix) parameterize the spread of the probability around the modes, but does not change the localization of the modes. If as before, the prior probabilities for the model parameters are assigned to be uniform, the logarithm of the posterior results in:

$$\ln P = -\sum_{m=1}^{M}\sum_{r=1}^{R}\sum_{\omega}\left[\ln\Gamma(\omega) + \Gamma(\omega)^{-1}\left|z_{mr}(\omega) - \sum_{n}^{N} B_{mn}a_{nr}s_n(\omega)e^{-i\omega\tau_{nr}}\right|^2\right] + const. \quad (4.50)$$

The MAP solutions can be derived as before for the amplitudes, latencies and matrix **B**. However, the derivation of the optimal solution for the component waveforms, $s_n(\omega)$, presents a problem. The difficulty being that the partial derivative of the log of the posterior $\frac{\partial \ln P}{\partial s_j(\omega)}$ involves now a non-analytic term $s_j^*(\omega)$. A potential workaround to this problem would be to decompose the AERP into one or several components, and then to estimate $s_n(\omega)$ from the Fourier transform of n^{th} component.

Another formulation in the spectral domain, given the stationarity condition for the ongoing activity is the *Z*-domain representation (Lange, Pratt, & Inbar, 1997):

$$z_{mr}(z) = \sum_{n=1}^{N} B_{mn}A_{nr}(z)s_n(z) + H_m(z)\varepsilon_{mr}(z), \quad (4.51)$$

where $\eta_{mr}(z)$ relates now to the autoregressive process $H_m(z)\varepsilon_{mr}(z)$ obtained by modeling the pre-stimulus activity, $A_{nr}(z)$ is a filter correcting for amplitude and phase

shift (latency) of the event related components, $H_m(z)=[1-\sum_{p=1}^{P}\alpha_p z^{-p}]^{-1}$ with α_p representing the coefficients of a P^{th}-order auto-regressive process obtained by modeling the pre-stimulus recordings. The auto-regressive model can be employed as a whitening filter of the ongoing activity by pre-whitening the data, i.e. $[H(z)]^{-1}z_{mr}(z)$. As a result the variance of the ongoing activity may be largely reduced, improving the performance of the estimation of the event-related components. Once the ongoing activity is pre-whitened, the subsequent derivation should also be similar to the one performed previously for the time domain representation.

Unfortunately, as seen in the previous Chapters, the ongoing activity can be highly non-stationary during the transition from pre to post-event, making questionable the use of pre-event recordings to estimate parameters for pos-event quantities. A promising alternative to be tested and compared in the future is then to include the power spectrum as a parameter to be estimated:

$$p(\{s_n(\omega)\},\{a_{nr}\},\{\tau_{nr}\},\mathbf{B},\{\Gamma_m(\omega)\}|\{z_{mr}(\omega)\},I)=$$

$$\frac{p(\{z_{mr}(\omega)\}|\{s_n(\omega)\},\{a_{nr}\},\{\tau_{nr}\},\mathbf{B},\{\Gamma_m(\omega)\},I)\;p(\{s_n(\omega)\},\{a_{nr}\},\{\tau_{nr}\},\mathbf{B},\{\Gamma_m(\omega)\}|I)}{p(\{z_{mr}(\omega)\})} \tag{4.52}$$

4.7 Discussion

In this Chapter, a principled way of estimating single trial multi-component event-related responses was presented. The approach is very promising from a methodological

point of view. The Bayesian formulation offers a truly statistical modeling strategy in the sense that shortcomings of the solutions can promptly point to the necessary reformulation of the model and prior probabilities. This is different from approaching the problem with the mindset restricted to the assumptions present in particular solutions like ICA, PCA, and other pure maximum likelihood techniques (Bell & Sejnowski, 1995; Jaskowski & Verleger, 1999; Jaskowski & Verleger, 2000; Jung, et al., 1999; Lange, Pratt, & Inbar, 1997; Pham, Mocks, Kohler, & Gasser, 1987; Woody, 1967).

The derived heuristic algorithm seems to be especially well suited for the estimation of relatively slow and large amplitude event-related components. Fast components (e.g. evoked gamma bursts) still represent a very difficult challenge to the formulation here because of their low signal-to-noise ratio and the larger impact of trial-to-trial latency variability when recorded from broad local fields. Another important issue is the choice of the number of event-related components, especially in the case of single channel recordings. Although differential variability in latency and amplitude of different components can add significantly to the separation of the components, especially in the case of the multi-channel estimation (Shah, et al., 2001), we believe that the knowledge about the underlying physiology still is a strong requirement for the choice of the number of event-related components, and for the initial guesses of their relative position in time and ranges for their trial-to-trial variability in latencies and amplitudes.

From the results in this Chapter and Chapter 3 there is ample evidence that the ongoing activity can be highly non-stationary during the transition from the pre- to the post-event period. As a consequence, it is difficult to introduce knowledge about the

power spectrum of the ongoing activity based on power spectrum estimates of the pre-event period recordings, as in the frequency domain formulation of the probabilistic model. The alternative of including the power spectrum of the ongoing activity as a parameter to be estimate simultaneously with the phase-locked component's parameters, may also be problematic if the adopted MAP solutions result in maximum likelihood procedures. Maximum Likelihood Estimates (MLE) will lead in general to the minimization of the power of the ongoing activity and over fitting of the phase-locked components. This is true also for MLE in time domain models as employed here. Thus, careful consideration of the dependence of the derived algorithm on the signal-to-noise ratio is necessary to avoid spurious results originating from "locking" of the estimated waveforms onto the ongoing activity. It should be kept in mind also that the introduction of knowledge of the temporal correlations in the ongoing activity does not change the localization of the modes of the posterior probability, but only their width, i.e., the spread of the probability. In this way the MAP solutions should be the same by assuming either white ongoing activity or temporally correlated ongoing activity. Any possible improvement in estimation after introduction of the knowledge of power spectrum should result from changes in the error bars for the estimators. That is because the error bars of the estimators are related to the width of the modes.

Comparing the derivation of this approach to the Bayesian derivation of Bell and Sejnowski's ICA algorithm (Knuth, 1997; Knuth, Truccolo, Bressler, & Ding, submitted) provides some insights into the relationships between these two techniques. While both algorithms assume that the signal mixing is linear, stationary and instantaneous, the

present approach does not require independence of the sources. This is key in neuroelectrophysiology as the dynamical interactions between neural source generators are a matter of great scientific interest. In addition, this algorithm accommodates noise as well as the possibility of the number of sources being different from the number of detectors. Finally, by defining the source model to explicitly allow for the potential variability of the source activity in individual trials we make this additional information available to aid in source identification. The power of this differential variability is being currently investigated in the context of multi-channel mulit-component estimation, where we believe it will be more effective. In this algorithm, we chose as a model of the source waveforms a set of discrete points describing the waveform amplitude at regular intervals. This source model is typically used in Blind Source Separation (BSS) and ICA applications, where the prior probability of the source amplitudes have been given by super- or sub-Gaussian probability densities. It is important to note, for this algorithm and others, that other source models are possible and in many cases desirable. These models could be continuous in nature (especially in the case of continuous latency shifts) such as linear-piecewise or cubic spline models, or could be dynamical in nature such as linear autoregressive moving average (ARMA) models. Finally, as described in previous works (Knuth, 1998; Knuth, 1999; Knuth & Vaughan, 1999), there are often cases where the experimenter has knowledge about the forward problem, which describes the propagation of the signals to the detectors. In such situations, one can incorporate information about the forward problem into the algorithm along with information about the geometry of the detector array by deriving appropriate prior probabilities for the coupling (or mixing)

matrix. In situations where the source locations are of interest, abandoning the coupling matrix in favor of a more detailed model of the source positions and orientations may be more fruitful. By modeling the source locations in addition to the source waveforms, one can easily design an algorithm that simultaneously performs source separation and localization (Knuth & Vaughan, 1999).

Further developments should include a detailed analysis of the convergence of the iterative algorithm as well as of the statistical properties of the estimators. As for the latter, the analysis should provide an explicit formulation of the error bar of the estimators and show how their variance approaches the Cramer-Rao lower bound. When working with alternative probabilistic models, the problem of model selection is again in place. Here, we have not considered alternative models to the generative model and so the issue of model selection was not addressed. This is a direction for the further development of estimation framework. An introduction to the problem of model selection from a Bayesian perspective can be found in Sivia (Sivia, 1996) and MacKay (Mackay, 1992). Also, although we have outlined the foundations for the MAP solution for multicomponents in the frequency domain, it has not been tested and compared to the performance of the time domain algorithm.

Regarding the problem of identifying the sources of the observed event-related modulations of interdependence measures computed on the original residual time series, the results presented in this Chapter indicates that these modulation were indeed originated from the trial-to-trial variability of the phase-locked components. The most definite evidence came from the great reduction in the 12 Hz coherence maximum values

when computed the new residuals were used for computation. Another important and related aspect is that the variance of the ongoing activity is nonstationary. This can be seen clearly in the case of the parietal channel example, and to a less extent in the striate channel recordings. The decrease in the variance of the parietal ongoing activity is probably related to the interruption of the ~ 22 Hz pre-stimulus oscillations just after the initial sensory activation. This decrease in both striate and parietal channel may also point out to the possibility that the sensory activation results in a strong synchronized event phase-locked activity that captures most of the dynamics, leading to a large scale re-organization of the ongoing activity.

Chapter 5

General Discussion

Two main possible sources of fast event-related transients in interdependence measures, namely the nonlinear gain modulation effect and the trial-to-trial variability of event phase-locked potentials, were analyzed. Very characteristic effects were revealed and tested on a particular data set consisting of local field potentials recorded intracortically from monkeys performing a visuo-motor pattern discrimination task. The results shown in Chapter 3 and 4 indicate that the observed event-related transient increases in variance, power, cross-correlation and 12 Hz coherence occurring about 100 ms after stimulus onset arose mainly from the combination of two factors: (1) around this time, there is a widespread activation originating perhaps from the striate visual cortex, resulting in event phase-locked signals in many different cortical sites; and (2) there is a

significant trial-to-trial variability in the amplitude and latency of these phase-locked signals. As a consequence, the variability of the transient phase locked signals leads to a transient increase in the variability of the residual time series. And, because of the widespread activation, there is also a widespread co-variation in amplitude, and possibly in latency too, among the event-phase locked signals from different recording sites. This co-variation, together with the signal-to-noise ratio effect, accounts for the observed transient increases in cross-correlation and 12 Hz coherence. Thus, the observed transients do not represent major event-related changes in functional connectivity, i.e. effective connectivity, among the many cortical sites involved in the episode.

There is, however, a large re-organization of the ongoing activity at the local level in many of the measured cortical sites. Our prototypical example has been the changes in the parietal site activity where a dramatic decrease in the predominant pre-stimulus ongoing ~ 22 Hz oscillation was observed. Similar events were also observed in the recorded somatosensory and motor areas. For cortical sites in the ventral pathway of the visual system, there was significant coherence around 15 Hz. This coherence pattern was also terminated after the initial sensory activation. The transition from pre-stimulus networks to the post-stimulus networks involves more complex mechanisms than that accounted by the proposed nonlinear gain modulation or by the trial-to-trial nonstationary of the evoked responses. This phenomenon also suggests that the evoked activity captures most of the dynamics, since most of the variability can be explained by the trial-to-trial variations of the event-phase locked components.

The reorganization of the ongoing activity during initial sensory activation indicates the oversimplification nature of our initial generative model where the linear superposition was assumed. But, a more appropriate model is yet unclear. The specific characteristics of this reorganization will probably depend highly on the properties of the local neuronal populations and the particular time scale of their activity.

It could be that the observed variations in amplitude and latency originated actually from multiplicative effects involving the ongoing and evoked activities. A preliminary result in Chapter 4 tells us that if there is a multiplicative effect, it does not express itself in terms of a linear correlation between levels of pre-stimulus power and evoked response amplitude and latency.

The theoretical and experimental results present here led to the analysis framework that was outlined in Chapter 1 and put in practice in Chapter 4. In summary, the proposed framework addresses three main challenges when looking at spatio-temporal interdependence patterns in event-related local field potential data: (a) nonstationarity of the brain activity stemming from rapid transitions between different processing stages occurring on the millisecond time scale; (b) latent or hidden variable problems stemming from the existence of mixed signals, including event phase-locked components, event non-phase-locked induced ongoing activities, event-unrelated ongoing activities and measurement noise; and (c) the trial-to-trial variability of event phase-locked signals.

To deal with these challenges, three main aspects were fundamental: (a) the proposed generative model; (b) the Bayesian inference approach; and (c) the parametric modeling of the ongoing activity by applying stochastic linear models. As mentioned

before, ideally the generation of phase-locked signals and ongoing activities should be modeled as the result of a single dynamical system. Improvements in this direction will be possible as more detailed experimental data are made available. That may also lead to different decompositions of the recorded signals, or at least, lead to better characterization of the physiological counterpart of the decomposed signals. For example, it may be that phase-locked signals recorded in primary sensory cortices related more to thalamic afferents, while in higher-level cortical areas they relate to different processes. Advancements in the Bayesian inference framework, particularly important when dealing with low signal-to-noise ratio recordings, are expected to result from the use of more informative prior probability models, and the application of better optimization techniques. Again, more informative priors will depend on new experimental data providing detailed biophysical properties of the neuronal generators. The need of parametric models of ongoing activities for the statistical analysis of interdependence was motivated by the fact that, given the nonstationarity of the data, short time windows were required. Further developments upon the choice of linear auto-regressive models can be made, obviously, by allowing nonlinearities, and also, by contemplating the possibility of higher-order statistical properties in the data, in other words, nongaussianity. Despite these alternatives, it should be mentioned that there are some evidence already suggesting that the inclusion of nonlinear terms in auto-regressive models may not result in better models. In cases other than, for example, time series recorded during epileptic states, quadratic and higher terms tend to be not significantly different from zero (N. Schiff, personal communication to my advisor Dr. M. Ding), suggesting that linear stochastic

models may be able to capture most of statistical properties of the relevant neural dynamics. Regarding the identification of higher-order statistics, a recent attempt in this direction has been done with the application of Volterra kernels (Friston, 2000). There is a potential problem with this approach though: the same trial-to-trial nonstationarities investigated in this dissertation may lead to nonzero higher terms in the Volterra kernels, which may then be mistakenly identified as the expression of nongaussianity or nonlinearity. The identification of higher-order statistics, under the constraint of the need of parametric modeling, may be solved by the development of nongaussian autoregressive models, i.e, models with 2^{nd} order temporally uncorrelated noise excitations but with nonzero higher order correlations (Nikias & Petropulu, 1993).

Appendices

For the sake of completeness, an extensive stability and bifurcation analysis of the neuronal model employed to study the nonlinear gain modulation effect is presented. In appendix A, the stability analysis leads to explicit constraints on the coupling strengths for the equilibrium point in this type of oscillatory networks. The stability analysis is in itself of interest since it provides original and significant results on the coupling constraints for the stability of both symmetrically and asymmetrically coupled neural networks of arbitrary size. In appendix B, the bifurcation analysis of the network's dynamics clarifies the relation between the sigmoid function and the observed bifurcation scenarios.

6.1 Appendix A: Stability Conditions for Equilibrium Points in Oscillatory Neural Networks

For oscillatory neural networks the stability of the equilibrium point is an important property for many neural computational applications. Aided by results from nonnegative matrices we demonstrate, using two simple models of coupled neural populations of arbitrary size, a general methodology that can yield explicit bounds on the individual coupling strengths to ensure the stability of the equilibrium point. Previous work on stability constraints have concentrated mainly on global stability in recurrent networks of the Hopfield type (Atiya, 1988; Chen & Amari, 2001; Cohen & Grossberg, 1983; Fang & Kincaid, 1996; Forti & Tsei, 1995; Guan, Chen, & Qin, 2000; Guez, Protopopsecu, & Barhen, 1988; Hirsch, 1989; Kelly, 1990; Liang & Wu, 1998; Matsuoka, 1991; Michel, Farrell, & Porod, 1989; Sugawara, Harao, & Noguchi, 1983; Yang & Dillon, 1994) and produced bounds on complicated functions of the coupling strengths. In this appendix we consider an approach that can lead to stability bounds on the individual coupling strengths in recurrent networks that can also admit more complex local dynamics. Two explicit models of coupled neural populations will be used to illustrate the approach.

6.1.1 Models

Our model's general topology is shown in Figure 6-1. The basic unit in the model is a neural population consisting of either excitatory x or inhibitory y cells (Freeman, 1975; Wilson & Cowan, 1972). The functional unit in the network is a cortical column consisting of mutually coupled excitatory and inhibitory populations. (By functional unit

we mean the simplest unit that is capable of oscillation.) The columns are then coupled through mutually excitatory interactions to form the network.

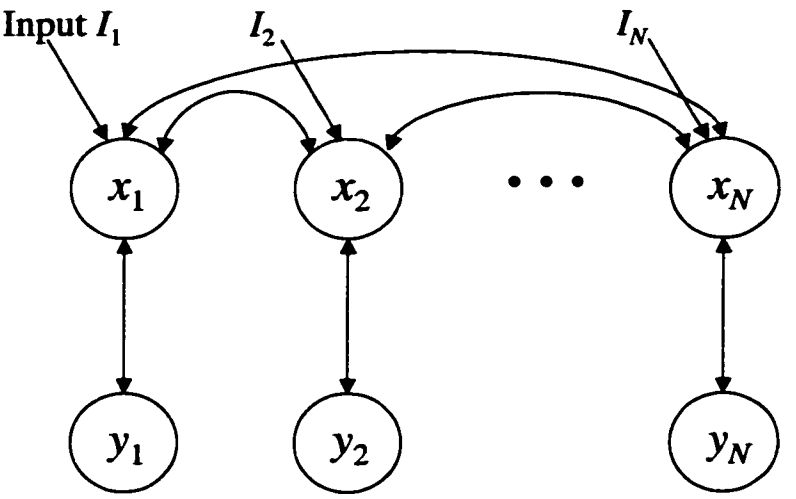

Figure 6-1: Model schematics

The two mathematical models considered here differ in their representation of the individual population. In the first model, henceforth referred to as Model A, which is similar to that considered in (Li & Hopfield, 1989; Wilson & Cowan, 1972) for modeling the olfactory bulb, each column is described by a system of two first order differential equations:

$$\begin{aligned} \frac{dx}{dt} + ax &= -k_{ei} Q(y, Q_m) I(t) \\ \frac{dy}{dt} + by &= k_{ie} Q(x, Q_m) \end{aligned} \qquad (6.1)$$

Here x, y represent the local field potentials of the excitatory and inhibitory populations, respectively, and I is the input. The constants a and b are the damping constants. The parameter $k_{ie} > 0$ gives the coupling gain from the excitatory (x) to the inhibitory (y) population whereas $k_{ei} > 0$ represents the strength of the reciprocal coupling. Further, $Q(v, Q_m)$ is taken to be a sigmoid function (Eeckman & Freeman, 1991a; Freeman, 1979):

$$\begin{aligned} Q(v, Q_m) &= Q_m[1 - \exp(-e^v - 1)/Q_m], \text{ if } v > -u_o \\ &= -1 \quad \text{if } v \le -u_o \end{aligned}, \tag{6.2}$$

where $u_o = -\ln[1 + \ln(1 + 1/Q_m)]$. It represents pulse densities converted from x controlled by a modulatory parameter Q_m. Note that $Q(0, Q_m) = 0$ and $Q'(0, Q_m) = 1$. (We note that the specific form of the sigmoid function is not important in our analysis. The function in Equation (6.2) is used for concreteness. The N columns are coupled together in the following fashion:

$$\begin{aligned} \frac{dx_n}{dt} + ax_n &= -k_{ei} Q(y_n, Q_m) + \frac{1}{N-1} \sum_{p=1}^{N} k_{np} Q(x_p, Q_m) + I_n(t) \\ \frac{dy_n}{dt} + by_n &= k_{ie} Q(x_n, Q_m) \end{aligned}, \tag{6.3}$$

where the columns are indexed by $n = 1,2,\ldots,N$ and the coupling strength k_{np} is the gain from the excitatory population of column p to the excitatory population of column n with $k_{nn} = 0 \ \forall n$.

The first order differential equation representation of a neural population in Model A has the property that its impulse response has an instantaneous rise phase. The benefit is that the resulting network is easier to analyze and can be cast in the format of a Hopfield network. In the second model, henceforth referred as Model B, the neural

population is a second differential equation possessing a finite rise and decay impulse response. Each individual column is described by a system of two second order differential equations (Freeman, 1992):

$$\begin{aligned} \frac{d^2x}{dt^2}+(a+b)\frac{dx}{dt}+abx &= -k_{ei}Q(y,Q_m)+I(t) \\ \frac{d^2y}{dt^2}+(a+b)\frac{dy}{dt}+aby &= k_{ie}Q(x,Q_m) \end{aligned} \quad . \tag{6.4}$$

For simplicity, we have assumed that the inhibitory and excitatory populations have identical rate constants. All the parameters have the same meaning as in Model A. The N column equations are given by:

$$\begin{aligned} \frac{d^2x_n}{dt^2}+(a+b)\frac{dx_n}{dt}+abx_n &= -k_{ei}Q(y_n,Q_m)+\frac{1}{N-1}\sum_{p=1}^{N}k_{np}Q(x_p,Q_m)+I(t) \\ \frac{d^2y_n}{dt^2}+(a+b)\frac{dy_n}{dt}+aby_n &= k_{ie}Q(x_n,Q_m) \end{aligned} \quad , \tag{6.5}$$

where as before $k_{nn}=0\ \forall n$. Model B is the same model used in Chapter 2.

6.1.2 Local Stability Analysis of Model A

First, we consider a single column given in Equation (6.1). When the input I is zero, it is easily seen that $x=0, y=0$ is an equilibrium point of the above dynamical system. In order to study its stability properties, we consider the Jacobian matrix **A** obtained by linearizing the above equations around the origin:

$$\mathbf{A}=\begin{pmatrix} -a & -k_{ei} \\ k_{ie} & -b \end{pmatrix}. \tag{6.6}$$

Here we have used the fact that $Q'(0,Q_m)=1$. For stability of the origin, the real parts of all eigenvalues of **A** should be less than zero. This requires

$$(a+b)>0,\ (ab+k_{ie}k_{ei})>0. \tag{6.7}$$

Since $a,b,k_{ei},k_{ie}>0$, these conditions are automatically satisfied. Hence the origin is a stable equilibrium point.

Next, we consider the stability of a network of columns described by Equation (6.3). As before, the origin is an equilibrium point for this network when input is absent. We study its stability by linearizing the above equations around the origin. For this, it is convenient to define matrices **S**, **K** and **C** as follows:

$$\mathbf{S}=\begin{pmatrix}\delta x_1 & \delta x_2 & \cdots & \delta x_N\\ \delta y_1 & \delta y_1 & \cdots & \delta y_N\end{pmatrix}, \tag{6.8}$$

$$\mathbf{K}=\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}, \tag{6.9}$$

and

$$[\mathbf{C}]_{np}=\frac{k_{np}}{N-1}, \tag{6.10}$$

if $n\neq p$ and zero otherwise. Here $\delta x_i,\delta y_i$ are deviations from the origin of x_i,y_i respectively. **S** is a 2 x N matrix containing the linearized variables of the model, **K** specifies the connectivity pattern and the N x N contains the coupling values between the columns. Upon linearizing Equation (6.3) we obtain the following matrix equation:

$$\frac{d\mathbf{S}}{dt}=\mathbf{AS}+\mathbf{KSC}. \tag{6.11}$$

We now consider two cases: a) The coupling matrix **C** is symmetric and b) **C** is not symmetric.

6.1.2.1 Symmetric Coupling Matrix

If **C** is symmetric, it can be written as

$$\mathbf{C} = \mathbf{B}\Lambda\mathbf{B}^{-1}, \tag{6.12}$$

where **B** and Λ are eigenvector and eigenvalue matrices of **C**, respectively. Let $\mathbf{e}(n)$ be one of the eigenvalues of **B** and $\lambda(n)$ its associated eigenvalue. Acting Equation (6.11) on $\mathbf{e}(n)$ we obtain the following eigenmode equation:

$$\frac{d\mathbf{Se}(n)}{dt} = \mathbf{ASe}(n) + \mathbf{K}\lambda(n)\mathbf{Se}(n). \tag{6.13}$$

Let $\mathbf{u}(n) = \mathbf{Se}(n)$. Then the above equation reduces to

$$\frac{d\mathbf{u}(n)}{dt} = [\mathbf{A} + \mathbf{K}\lambda(n)]\mathbf{u}(n). \tag{6.14}$$

Note that all the eigenvalues $\lambda(n)$ are real since **C** is assumed to be symmetric. Using the above procedure, we have reduced the study of a N-dimensional system to the study of a 2-dimensional system.

We now return to the question of stability of the origin. It will be stable if

$$\text{Re } \alpha_i(n) < 0, \ \forall i, n, \tag{6.15}$$

where $\alpha_i(n)$'s are the eigenvalues of the matrix $\mathbf{A} + \mathbf{K}\lambda(n)$. Suppressing the indices n, i for notational simplicity, the characteristic polynomial of this matrix is given by

$$\alpha^2 + \alpha(a + b - \lambda) + (k_{ei}k_{ie} + ab - b\lambda). \tag{6.16}$$

The range of parameter values that gives Re $\alpha < 0$ can be determined by applying the Lienard-Chipart criterion (see Appendix Section 6.1.5). Since the coefficient of α^2 is greater than zero, the Lienard-Chipart criterion reduces to

$$(a + b - \lambda) > 0, \ (k_{ei}k_{ie} + ab - b\lambda) > 0. \tag{6.17}$$

This implies that

$$\lambda < (a+b),\ \lambda < \frac{k_{ei}k_{ie}}{b} + a. \tag{6.18}$$

Both these inequalities are simultaneously satisfied if

$$\lambda < K^{\max} = \min\{(a+b), \frac{k_{ei}k_{ie}}{b} + a\}. \tag{6.19}$$

where $K^{\max}$ is obviously positive. But this has to be satisfied for every eigenvalue λ of **C**. This is ensured if (denoting the largest eigenvalue of **C** by $\lambda^{\max}(\mathbf{C})$)

$$\lambda^{\max}(\mathbf{C}) < K^{\max}. \tag{6.20}$$

We now translate the upper bound on the maximum eigenvalue of **C** to an upper bound on the individual coupling coefficients constituting **C** using the theory of nonnegative matrices (Horn & Johnson, 1990). Note that **C** is a nonnegative matrix since all its entries are nonnegative. We introduce a N x N matrix **U** whose matrix entries are given by

$$[\mathbf{U}]_{ij} = \frac{K^{\max}}{N},\ i,j = 1,2,\ldots,N. \tag{6.21}$$

Since $K^{\max} > 0$, all entries are greater than zero and **U** is a positive matrix. Further, $\lambda^{\max}(\mathbf{U}) = K^{\max}$. If we impose the following condition on the coupling coefficients k_{np}:

$$0 \le k_{np} < \frac{N-1}{N} K^{\max},\ n \ne p, \tag{6.22}$$

we see that $[\mathbf{C}]_{ij} - [\mathbf{U}]_{ij} < 0$ for i,j. Hence $\mathbf{C} < \mathbf{U}$. Then, from nonnegative matrix theory (Horn & Johnson, 1990), we get

$$\lambda^{\max}(\mathbf{C}) < \lambda^{\max}(\mathbf{U}) = K^{\max}. \tag{6.23}$$

But this is the required condition for stability of origin. To summarize, origin is a stable fixed point for the neural network given in Equation (6.3) with symmetric coupling if the coupling strengths obey the constraints given in Equation (6.22), where K^{max} is given in Equation (6.19) and is only a function of intra-column parameters.

6.1.2.2 Non Symmetric Coupling Matrix

Next, we consider the case when the coupling matrix **C** is not symmetric. In this case, it can be written as

$$\mathbf{C} = \mathbf{BJB}^{-1}, \tag{6.24}$$

where **J** is the Jordan canonical form of **C**. Corresponding to each Jordan block in **J** there is only one distinct eigenvalue $\lambda(n)$ and its associated eigenvector $\mathbf{e}(n)$ (Horn & Johnson, 1990). Acting Equation (6.11) on $\mathbf{e}(n)$ we get

$$\frac{d\mathbf{u}(n)}{dt} = [\mathbf{A} + \mathbf{K}\lambda(n)]\mathbf{u}(n). \tag{6.25}$$

The origin will be stable if

$$\text{Re } \alpha_i(n) < 0, \ \forall i, n, \tag{6.26}$$

where $\alpha_i(n)$'s are the eigenvalues of the Jacobian matrix $\mathbf{A} + \mathbf{K}\lambda(n)$. Suppressing the indices n,i for notational simplicity, the characteristic polynomial of this Jacobian matrix is given by

$$\alpha^2 + \alpha(a + b - \lambda) + (k_{ei}k_{ie} + ab - b\lambda). \tag{6.27}$$

Since **C** is not assumed to be symmetric, λ can be a complex number. Writing it as:

$$\lambda = \lambda_R + i\lambda_I, \tag{6.28}$$

the characteristic polynomial becomes

$$f(\alpha) = \alpha^2 + \alpha(a+b-\lambda_{_R} - i\lambda_{_I}) + (k_{ei}k_{ie} + ab - b\lambda_R - ib\lambda_I). \qquad (6.29)$$

The range of parameter values that gives Re $\alpha < 0$ can be determined by applying the generalized Routh-Hurwitz criterion (see Section 6.1.6). Following this procedure, consider $-if(i\alpha)$:

$$-if(i\alpha) = i\alpha^2 + \alpha(a+b-\lambda_R) - i\alpha\lambda_I - i(k_{ei}k_{ie} + ab - b\lambda_R) - b\lambda_I. \qquad (6.30)$$

This has to be put into the following standard form:

$$-if(i\alpha) = b_0\alpha^2 + b_1\alpha + b_2 + i[a_0\alpha^2 + a_1\alpha + \alpha_2]. \qquad (6.31)$$

Comparing the two equations we get

$$a_0 = 1,\ a_1 = -\lambda_I,\ a_2 = -(k_{ei}k_{ie} + ab - b\lambda_R), \qquad (6.32)$$

$$b_0 = 0,\ b_1 = (a+b-\lambda_R),\ b_2 = -b\lambda_I. \qquad (6.33)$$

Applying the generalized Routh-Hurwitz criterion, we have Re $\alpha < 0$ if the following two conditions are met:

$$\nabla_2 = \begin{vmatrix} 1 & -\lambda_I \\ 0 & (a+b-\lambda_R) \end{vmatrix} > 0, \qquad (6.34)$$

and

$$\nabla_2 = \begin{vmatrix} 1 & -\lambda_I & -(k_{ei}k_{ie} + ab - b\lambda_R) & 0 \\ 0 & (a+b-\lambda_R) & -b\lambda_I & 0 \\ 0 & 1 & -\lambda_I & -(k_{ei}k_{ie} + ab - b\lambda_R) \\ 0 & 0 & (a+b-\lambda_R) & -b\lambda_I \end{vmatrix} > 0. \qquad (6.35)$$

Evaluating the above determinants and simplifying, we get

$$(a+b-\lambda_R) > 0, \qquad (6.36)$$

$$(k_{ei}k_{ie} + ab - b\lambda_R)(a+b-\lambda_R)^2 - b\lambda_I^2(\lambda_R - a) > 0. \tag{6.37}$$

If **C** is symmetric, $\lambda_I = 0$ for all eigenvalues and the above inequalities reduce to stability condition given in Equation (6.19), as expected. It is clear that if $\lambda_R < a$, both inequalities are satisfied irrespective of the value of λ_I. Thus, a sufficient (but not necessary) condition for stability of the origin is given as follows. For every eigenvalue λ of the coupling matrix **C**, we require

$$\lambda_R < a. \tag{6.38}$$

To proceed further, consider the spectral radius $\rho(\mathbf{C})$ of **C**:

$$\rho(\mathbf{C}) = \max\{|\lambda|, \lambda \text{ is an eigenvalue of } \mathbf{C}\}. \tag{6.39}$$

Since **C** is a nonnegative matrix, it turns out (Horn & Johnson, 1990) that $\rho(\mathbf{C})$ itself is the largest eigenvalue of **C**. That is,

$$\lambda^{\max} = \rho(\mathbf{C}). \tag{6.40}$$

Further, from the definition of $\rho(\mathbf{C})$,

$$\mathrm{Re}\,\lambda \le \lambda^{\max} = \rho(\mathbf{C}) \;\forall\, \lambda. \tag{6.41}$$

Therefore, the stability condition given in Equation (6.38) reduces to

$$\rho(\mathbf{C}) < a. \tag{6.42}$$

This stability condition has a nice physical interpretation. Note that a is nothing but a damping constant for the equation in x variable [cf. Equation (6.3)] whereas the coupling coefficients which constitute **C** represent the strength of the intercolumnal excitation. Therefore, if the "size" of **C** is less than a, the system will be damped and origin will be stable (as $|x|$ decreases, the function $Q(x, Q_m)$ also decreases to zero and therefore y is

also damped). The size of **C** is measured by its matrix norm and the natural norm in this case is the spectral norm. But the spectral norm is identical to the spectral radius (Horn & Johnson, 1990). Hence, from physical considerations we expect the origin to be stable if the spectral radius of **C** is less than a. This is nothing but the stability condition derived in Equation (6.42).

If a necessary and sufficient condition for stability is required, one can solve the above inequalities numerically for the given set of parameters and obtain the region in complex space where the inequalities are satisfied. But this region would have complicated boundaries for $a < \lambda_R < \min\{(a+b), \frac{k_{ei}k_{ie}}{b} + a\}$ whereas it is simple for the case $\lambda_R < a$ considered above. Note that the upper bound on λ_R arises from the fact that λ_I^2 has to be nonnegative [cf. Equation (6.37)].

We now translate the sufficient condition for stability given in Equation (6.42) to an upper bound on the coupling coefficients. We follow the same procedure as in the symmetric case. We introduce a N x N matrix **U** whose matrix entries are given by

$$[\mathbf{U}]_{ij} = \frac{a}{N},\ i,j = 1,2,\ldots,N. \tag{6.43}$$

Since all entries are greater than zero, **U** is a positive matrix. Further, the spectral radius of **U** is given by $\rho(\mathbf{U}) = a$. If we impose the following condition on the coupling coefficients k_{np}:

$$0 < k_{np} < a\frac{N-1}{N},\ n \neq p, \tag{6.44}$$

we see that $[\mathbf{C}]_{ij} - [\mathbf{U}]_{ij} < 0$ for all i,j. Hence $\mathbf{C} < \mathbf{U}$. Then, from nonnegative matrix theory (Horn & Johnson, 1990), we get

$$\rho(\mathbf{C}) < \rho(\mathbf{U}) = a. \tag{6.45}$$

But this is the required condition for stability of origin [cf. Equation (6.42)]. To summarize, origin is a stable fixed point for the neural network given in Equation (6.3) with nonsymmetric coupling if the coupling strengths obey the constraints given in Equation (6.44).

6.1.3 Stability Analysis of Model B

In this section, we investigate the stability of Model B in detail. First, we consider the stability of the single column equations given in Equation (6.4). When the input I is zero, it is again seen that the origin $x = 0, y = 0$ is an equilibrium point of the above dynamical system. In order to study its stability properties, we first convert the above second order differential equations to the following system of first order differential equations (with $I = 0$):

$$\begin{aligned}
\frac{dz_1}{dt} &= z_2 \\
\frac{dz_2}{dt} &= -(a+b)z_2 - abz_1 - k_{ei}Q(z_3, Q_m) \\
\frac{dz_3}{dt} &= z_4 \\
\frac{dz_4}{dt} &= -(a+b)z_4 - abz_3 + k_{ie}Q(z_1, Q_m)
\end{aligned}, \tag{6.46}$$

where

$$z_1 = x,\ z_2 = \frac{dx}{dt},\ z_3 = y,\ z_4 = \frac{dy}{dt}. \tag{6.47}$$

Consider the Jacobian matrix **A** obtained by linearizing the above equations around the origin:

$$\mathbf{A}=\begin{pmatrix} 0 & 1 & 0 & 0 \\ -ab & -(a+b) & -k_{ei} & 0 \\ 0 & 0 & 0 & 1 \\ k_{ie} & 0 & -ab & -(a+b) \end{pmatrix}. \tag{6.48}$$

Here we have used the fact that $Q'(0,Q_m)=1$. For stability of the origin, the real parts of all eigenvalues of **A** should be less than zero. The eigenvalues are determined from the characteristic equation:

$$\lambda^4+2(a+b)\lambda^3+(a^2+4ab+b^2)\lambda^2+2(a^2b+ab^2)\lambda+k_{ei}k_{ie}+a^2b^2=0. \tag{6.49}$$

Applying the Lienard-Chipart criterion, the real parts of all eigenvalues are negative if the following inequalities be satisfied:

$$\begin{aligned} a^2b^2+k_{ei}k_{ie}&>0 \\ 2a^2b+2ab^2&>0 \\ 2a+2b&>0 \\ -4(a+b)^2k_{ei}k_{ie}+4(a+b)^4&>0 \end{aligned} \tag{6.50}$$

Since $a,b,k_{ei},k_{ie}>0$, the first three inequalities are automatically satisfied. After simplification, the last inequality can be written as:

$$k_{ei}k_{ie}<ab(a+b)^2. \tag{6.51}$$

To summarize, the origin is stable for the single column equations if the above condition is satisfied. Henceforth, we will assume that this is true. Next, we consider the stability of a network of columns given in Equation (6.5). As usual, we study the stability

of the origin (which is an equilibrium point) by linearizing the above equations around it. For this, it is convenient to define matrices **S**, **K** and **C** as follows:

$$\mathbf{S}=\begin{pmatrix} \delta x_1 & \delta x_2 & \cdots & \delta x_N \\ \frac{d\delta x_1}{dt} & \frac{d\delta x_2}{dt} & \cdots & \frac{d\delta x_N}{dt} \\ \delta y_1 & \delta y_2 & \cdots & \delta y_N \\ \frac{d\delta y_1}{dt} & \frac{d\delta y_2}{dt} & \cdots & \frac{d\delta y_N}{dt} \end{pmatrix}, \tag{6.52}$$

$$\mathbf{K}=\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{6.53}$$

and

$$[\mathbf{C}]_{np}=\frac{k_{np}}{N-1}, \tag{6.54}$$

if $n \neq p$ and zero otherwise. Here $\delta x_i, \delta y_i$ are deviations from the origin of x_i, y_i respectively. Further, **S** is a 4 x N matrix containing the linearized variables of the model, **K** specifies the connectivity pattern and the N x N matrix **C** contains the coupling values between the columns.

Upon linearizing Equation (6.5) we obtain the following matrix equation:

$$\frac{d\mathbf{S}}{dt}=\mathbf{AS}+\mathbf{KSC}. \tag{6.55}$$

We now assume that the coupling matrix is symmetric. Then all the eigenvalues of **C** are real. Following the same procedure as given in Section 6.1.2, we can reduce the above equation to a set of decoupled eigenmode equations:

$$\frac{d\mathbf{u}(n)}{dt}=[\mathbf{A}+\mathbf{K}\lambda(n)]\mathbf{u}(n). \tag{6.56}$$

As before, the origin will be stable if

$$\text{Re } \alpha_i(n) < 0, \ \forall i,n, \tag{6.57}$$

where $\alpha_i(n)$'s are the eigenvalues of $\mathbf{A}+\mathbf{K}\lambda(n)$. Suppressing the indices n,i for notational simplicity, the characteristic polynomial of this matrix is given by

$$\begin{aligned}&\alpha^4 + 2(a+b)\alpha^3 + [(a+b)^2 + 2ab - \lambda]\alpha^2 + [2ab(a+b) - \lambda(a+b)]\alpha + \\ &[a^2b^2 - ab\lambda + k_{ei}k_{ie}]\end{aligned}. \tag{6.58}$$

The range of parameter values which give Re $\alpha < 0$ can be determined by again applying the Lienard-Chipart criterion. We get the following inequalities:

$$\begin{aligned} a^2b^2 - ab\lambda + k_{ei}k_{ie} &> 0 \\ 2ab(a+b) - \lambda(a+b) &> 0 \\ 2(a+b) &> 0 \\ (a+b)^2\lambda^2 - 2(a+b)^4\lambda + 4(a+b)^2(a^3b + 2a^2b^2 + ab^3 - k_{ei}k_{ie}) &> 0 \end{aligned}. \tag{6.59}$$

Since a, b are positive, the third inequality is automatically satisfied. After simplification, the first two inequalities become:

$$\begin{aligned}&\lambda < \frac{k_{ei}k_{ie} + a^2b^2}{ab} \\ &\lambda < 2ab\end{aligned}, \tag{6.60}$$

The last inequality is of the form

$$a_1\lambda^2 - a_2\lambda + a_3 > 0, \tag{6.61}$$

where

$$a_1(a+b)^2, \ a_2 = 2(a+b)^4, \ a_3 = 4(a+b)^2[ab(a+b)^2 - k_{ei}k_{ie}]. \tag{6.62}$$

Note that a_1, a_2 are obviously positive. It turns out a_3 is also positive because of the local stability condition derived in Equation (6.51). The quadratic function $a_1\lambda^2 - a_2\lambda + a_3$ with

a_1, a_2, a_3 positive has a unique global minimum at $\lambda = a_2/2a_1$. Thus the minimum occurs at a positive value of λ. It is also seen that

$$a_2^2 - 4a_1a_3 = 4(a+b)^4\left[(a+b)^4 - 4[ab(a+b)^2 - k_{ei}k_{ie}]\right]. \tag{6.63}$$

This can be simplified as

$$a_2^2 - 4a_1a_3 = 4(a+b)^4\left[(a^2-b^2)^2 + 4k_{ei}k_{ie}\right], \tag{6.64}$$

which is positive since $k_{ei}k_{ie}$ is positive. Thus both the zeros of the quadratic function (we will denote them η_1 and η_2 with $\eta_1<\eta_2$) are real. Further, since $a_3 > 0$ and the global minimum occurs at a positive value, $\eta_2 > \eta_1 > 0$. Consequently, the last inequality is satisfied when $\lambda < \eta_1$ and $\lambda > \eta_2$ where

$$\eta_{1(2)} = (a+b)^2 - (+)\sqrt{(a+b)^4 - 4[ab(a+b)^2 - k_{ei}k_{ie}]}. \tag{6.65}$$

Note that η_1 is explicitly seen to be positive by applying Equation (6.51). Further, $\eta_2 > (a+b)^2 > 2ab$. Thus the inequality $\lambda > \eta_2 > 2ab$ is not possible given the stability condition $\lambda < 2ab$ derived earlier. Therefore the last inequality in Equation (6.59) reduces to $\lambda < \eta_1$.

Summarizing, we get the following set of stability conditions:

$$\begin{aligned} &\lambda < \frac{k_{ei}k_{ie} + a^2b^2}{ab} \\ &\lambda < 2ab \\ &\lambda < \eta_1 \end{aligned} \quad . \tag{6.66}$$

All these inequalities will be simultaneously satisfied if

$$\lambda < K^{\max},\ K^{\max} = \min\{\frac{k_{ei}k_{ie} + a^2b^2}{ab}, 2ab, \eta_1\} > 0. \tag{6.67}$$

Note that K^{max} is positive since all constituents we are minimizing over are positive. The above condition has to be satisfied for all eigenvalues of **C**. This is ensured if

$$\lambda^{max}(\mathbf{C}) < K^{max}, \tag{6.68}$$

where $\lambda^{max}(\mathbf{C})$ is the largest eigenvalue of **C**.

We now translate the above condition which guarantees the stability of the origin into bounds on the coupling coefficients. Note that **C** is a nonnegative matrix since all its entries are nonnegative. We introduce a N x N matrix **U** whose matrix entries are given by

$$[\mathbf{U}]_{ij} = \frac{K^{max}}{N}, \; i, j = 1,2,\ldots,N. \tag{6.69}$$

Since K^{max} is greater than zero, all entries are greater than zero and **U** is a positive matrix. Further, $\lambda^{max}(\mathbf{U}) = K^{max}$. If we impose the following condition on the coupling coefficients k_{np}:

$$0 \le k_{np} < \frac{N-1}{N} K^{max}, \; n \ne p, \tag{6.70}$$

we see that $[\mathbf{C}]_{ij} - [\mathbf{U}]_{ij} < 0$ for all i, j. Hence **C** < **U**. Then, from nonnegative matrix theory (Horn & Johnson, 1990), we get

$$\lambda^{max}(\mathbf{C}) < \lambda^{max}(\mathbf{U}) = K^{max}. \tag{6.71}$$

But this is the required condition for stability of origin. To summarize, origin is a stable equilibrium point for the neural network given in Equation (6.5) with symmetric coupling if the coupling strengths obey the constraints in Equation (6.70), where K^{max} is given in Equation (6.67) and is only a function of the parameters in a single column.

6.1.4 Relations to the Hopfield Neural Network Model and Global Stability

Guan et. al. (Guan, Chen, & Qin, 2000) study the following Hopfield neural network model:

$$C_i \frac{du_i}{dt} = \sum_{j=1}^{L} T_{ij} v_j - \frac{u_i}{R_i} + I_i, \; i = 1,2,\ldots,L, \tag{6.72}$$

$v_i = g_i(u_i), C_i, R_i > 0$, I_i gives the input and T_{ij} denote the coupling strengths (which need not be symmetric). For this system, they show that the equilibrium state is unique and globally asymptotically stable if the following conditions are met:

$$T_{ii} - \frac{1}{R_i M_i} + \sum_{j \neq i} |T_{ij}| < 0, \; i = 1,2,\ldots,L, \tag{6.73}$$

where $0 < g_i' \leq M_i$.

The above model reduces to Model A given in Equation (6.3) if we make the following identifications:

$$\begin{aligned} &L = 2N, \; \mathbf{u} = (x_1, x_2, \ldots, x_N, y_1, y_2, \ldots, y_N) \\ &C_i = 1 \; \forall i, \; g_i(u_i) = Q(u_i, Q_m) \\ &R_i = 1/a, \; i = 1,2,\ldots,N, \; R_i = 1/b, \; i = N+1, N+2, \ldots, 2N \end{aligned} \tag{6.74}$$

The coupling coefficients T_{ij} are identified as follows:

$$\begin{aligned} &T_{ij} = k_{ij}/(N-1), \; i,j = 1,2,\ldots,N \\ &T_{ij} = -k_{ei}, \; i = 1,2,\ldots,N, \; j = N+1, N+2, \ldots, 2N, \\ &T_{ij} = k_{ie}, \; i = N+1, N+2, \ldots, 2N, \; j = 1,2,\ldots,N \end{aligned} \tag{6.75}$$

and $T_{ij} = 0$ otherwise. Evaluating the maximum value of $g_i' = Q'(u_i, Q_m)$ which occurs at $u_i = \ln Q_m$, we get $M_i = Q_m \exp(-(Q_m - 1)/Q_m) \; \forall i$.

Using the above correspondence between the two models, we find that our equilibrium solution is unique and globally stable if the following conditions are met:

$$\frac{1}{N-1}\sum_{p\le n} k_{np} + k_{ei} < a\exp[(Q_m - 1)/Q_m]/Q_m,\ n = 1,2,\ldots,N. \tag{6.76}$$

$$k_{ie} < b\exp[(Q_m - 1)/Q_m]/Q_m. \tag{6.77}$$

Here we use the fact that $\exp[(Q_m - 1)/Q_m]/Q_m \le 1$.

Note that this condition imposes constraints on k_{ei} and k_{ei} which were not required for local stability. This is reasonable since having a unique global stable equilibrium point is more stringent than having a linearly stable equilibrium point required in the present work. Model B in Equation (6.5) can no longer be cast in a Hopfield framework but our approach is applicable to both models.

6.1.5 The Lienard-Chipart Criterion

In this and in the following section, we state the Lienard-Chipart and the generalized Routh-Hurwitz criteria. The statements are taken directly from Gantmacher (Gantmacher, 1964) and are given here for the sake of completeness.

Consider a real polynomial

$$f(z) = a_0 z^n + a_1 z^{n-1} + \ldots + a_n, \tag{6.78}$$

with $a_0 > 0$. Necessary and sufficient conditions for all the zeros of the polynomial to have negative real parts can be given in any *one* of the following forms (Gantmacher, 1964):

(1) $a_n > 0, a_{n-2} > 0, \ldots; \Delta_1 > 0, \Delta_3 > 0, \ldots,$

(2) $a_n > 0, a_{n-2} > 0, \ldots; \Delta_2 > 0, \Delta_4 > 0, \ldots,$

(3) $a_n > 0, a_{n-1} > 0, a_{n-3} > 0, \ldots; \Delta_1 > 0, \Delta_3 > 0, \ldots,$

(4) $a_n > 0, a_{n-1} > 0, a_{n-3} > 0, \ldots; \Delta_2 > 0, \Delta_4 > 0, \ldots.$

Here Δ_p is the Hurwitz determinant of order p given by the formula

$$\Delta_p = \begin{vmatrix} a_1 a_3 a_5 \ldots & & \\ a_0 a_2 a_4 \ldots & & \\ 0 a_1 a_3 \ldots & & \\ 0 a_0 a_2 \ldots & & \\ & \cdot & \\ & .a_p & \end{vmatrix}, \quad p = 1,2,\ldots,n, \tag{6.79}$$

where $a_k = 0$ for $k>n$. In the literature, the equivalent Routh-Hurwitz criterion is usually used. But the Lienard-Chipart is better since the number of determinants that have to be evaluated is half the number that have to be evaluated for the Routh-Hurwitz criterion. This leads to a simpler set of inequalities that need to be evaluated. In the main text, we use the third form of the Lienard-Chipart criterion given above.

6.1.6 The Generalized Routh-Hurwitz Criterion

Consider a polynomial $f(z)$ with complex coefficients. Suppose that

$$f(iz) = b_0 z^n + b_1 z^{n-1} + \ldots + b_n + i(a_0 z^n + a_1 z^{n-1} + \ldots + a_n), \tag{6.80}$$

where $a_0, a_1, \ldots, a_n, b_1, \ldots, b_n$ are real numbers. If the degree of $f(z)$ is n, then $b_0 + ia_0 \neq 0$. Without loss of generality, we may assume that $a_0 \neq 0$. Otherwise, we consider the polynomial $g(z) = -if(z)$ and repeat the analysis for this polynomial. Both $f(z)$ and $g(z)$ have the same set of zeros and so no information is lost. This is the case considered in the main text.

If $\nabla_{2p} \neq 0$, then all the zeros of $f(z)$ have negative real parts if

$$\nabla_2 > 0, \nabla_4 > 0, \ldots, \nabla_{2n} > 0, \tag{6.81}$$

where

$$\nabla_{2p} = \begin{vmatrix} a_0 & a_1 & \cdots & a_{2p-1} \\ b_0 & b_1 & \cdots & b_{2p-1} \\ 0 & a_0 & \cdots & a_{2p-2} \\ 0 & b_0 & \cdots & b_{2p-2} \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix}, \quad p = 1,2,\ldots,n, \tag{6.82}$$

where $a_k = b_k = 0$ for $k > n$. Note that the condition $\nabla_{2n} \neq 0$ would be satisfied for a generic set of parameter values. This is especially true in our case where a_k, b_k are functions of system parameters.

6.2 Appendix B: Bifurcation Scenarios

In this section, we investigate how the dynamics evolve as a function of input variation. Our main goal is to discover and characterize qualitatively different bifurcation scenarios. Only the model B from the Appendix A will be analyzed here. The important role played by the nonlinear sigmoid function in this case will be elucidated.

6.2.1 Single Column

6.2.1.1 Bifurcations Under Input

From (6.4) and for a given level of input I, the fixed point (x_0, y_0) can be found by solving the following equations:

$$x_0 = \frac{1}{ab}(I - k_{ei}Q(\frac{k_{ie}}{ab}Q(x_0.Q_m),Q_m)$$
$$y_0 = \frac{k_{ie}}{ab}Q(x_0.Q_m) \tag{6.83}$$

Recalling that the sigmoid function $Q(v,Q_m)$ is a monotonically increasing function of v, one can see, from Equation (6.83), that (1) x_0 and y_0 increase monotonically as I increases and (2) for a given level of input there is only one fixed point. By linearizing around (x_0, y_0) we obtain

$$\frac{d^2x}{dt^2} + (a+b)\frac{dx}{dt} + abx = -k_{ei}Q'(y_0,Q_m)$$
$$\frac{d^2y}{dt^2} + (a+b)\frac{dy}{dt} + aby = k_{ie}Q'(x_0,Q_m) \tag{6.84}$$

The characteristic equation and the eigenvalues are found to be

$$(\lambda+a)^2(\lambda+b)^2 + k_{ie}Q'(x_0,Q_m)k_{ei}\ Q'(y_0,Q_m) = 0, \tag{6.85}$$

and

$$\lambda_{1,2} = \frac{-(a+b) \pm \sqrt{(a+b)^2 - 4[ab \pm i\sqrt{k_{ie}Q'(x_0,Q_m)k_{ei}Q'(y_0,Q_m)}]}}{2}. \tag{6.86}$$

The oscillatory nature of the fixed point is apparent since the eigenvalues are always complex for nonzero mutual coupling. Since the sigmoid function has variable slope providing a nonlinear gain, depending on the position of (x_0, y_0), the terms $Q'(x_0,Q_m)$ and $Q'(y_0,Q_m)$ in the eigenvalue expression can vary their values such that the stability of (x_0, y_0) might change, i. e., (x_0, y_0) becomes unstable. Specifically, one can show that if

$$Q'(x_0, Q_m)Q'(y_0, Q_m) \geq \frac{ab(a+b)^2}{k_{ie}k_{ei}}, \tag{6.87}$$

then the fixed point will become unstable as input increases. The resulting bifurcation is a supercritical Hopf bifurcation by which a stable limit cycle is created. We note that the value of Q_m plays an important role for the condition in Equation (6.87) to be satisfied. Figure 6-2 illustrates the system's response as I increases for three different values of Q_m.

For all different values of Q_m, as I further increases, the nonlinear gain passes through its maximum value and begins to decrease, leading to the fixed point regaining its stability at a larger value of I. Another notable feature of the bifurcation diagram is that for large enough I, the fixed point position depends on I in a linear fashion. This can be understood through the following analysis. As I goes to infinity

$$Q(x_0, Q_m) \to Q_m \text{ and } Q(\frac{k_{ie}}{ab}Q(x_0, Q_m), Q_m) \to Q(\frac{k_{ie}}{ab}Q_m, Q_m). \tag{6.88}$$

From Equation (6.83) the fixed point position x_0 and y_0 in the upper saturation of the sigmoid as a function of I is solved to be:

$$\begin{aligned} x_0 &= \frac{1}{ab}(I - k_{ei}Q(\frac{k_{ie}}{ab}Q_m, Q_m)) \\ y_0 &= \frac{k_{ie}}{ab}Q_m \end{aligned} \tag{6.89}$$

These relations are in excellent agreement with the linearly increasing portion of the bifurcation diagram in Figure 6-2.

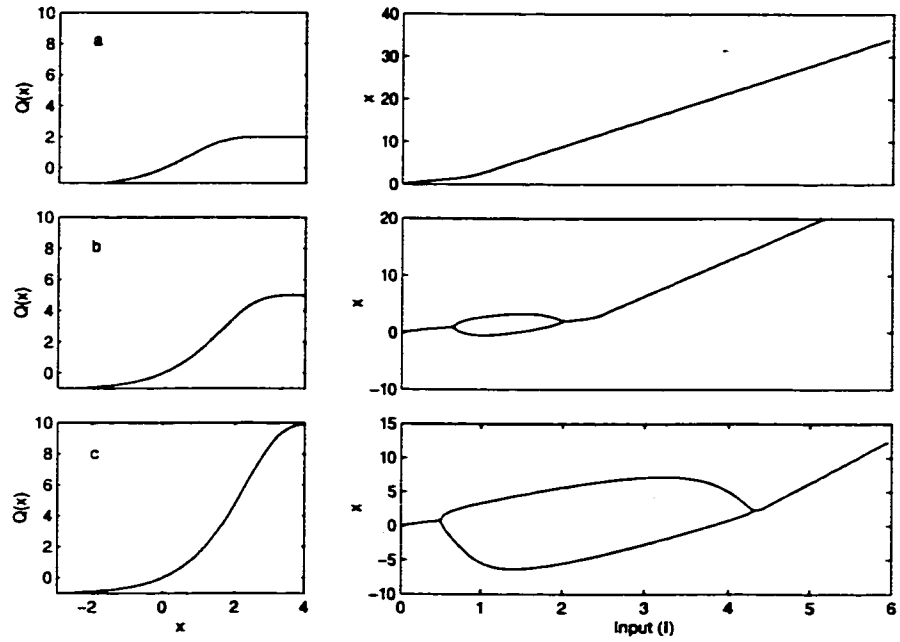

Figure 6-2: Single column bifurcation diagrams as function of the parameter Q_m. The bifurcation parameter is the input level I. The left plot is the sigmoid function, $Q(x)$, and on the right the respective bifurcation diagram for the excitatory population (x). For Q_m=2, the slope or gain provided by the sigmoid output function is not high enough to destabilize the fixed point. The Hopf bifurcation does not happen and the equilibrium point is moved to higher levels as input increases (top panel). In the middle panel, we choose Q_m=5 and the single column exhibits a Hopf bifurcation. The limit cycle is represented by its maximum and minimum amplitude (up and low branches of the "bubble"). Higher levels of Q_m make the Hopf bifurcation to occur earlier and the limit cycle is present for a larger interval of input levels. This is seen in the bottom panel, where Q_m=10. The other parameters are: $a = 220/s, b = 720/s, k_{ie} = 0.1$ and $k_{ei} = 0.4$.

6.2.1.2 Frequency and Phase

Depending on the input level the dynamics of a single column is either a damped oscillation or a sustained oscillation. The relative phase between the excitatory and inhibitory populations is about a quarter of a cycle (90°), with the inhibitory population lagging the excitatory one for the sustained oscillations. Below we give an estimate of the

frequency of these oscillations by a linear approximation. Assume that the input level is such that the fixed point is stable. Any nonzero initial condition will be attracted to the fixed point via a damped oscillation. From the eigenvalue equation (6.86), the frequency F of the damped oscillation is approximately:

$$F \approx \frac{\mathrm{Im}(\lambda)}{2\pi} 10^3 \mathrm{Hz}, \tag{6.90}$$

if the rate constants are given in milliseconds. Although Equation (6.90) is based on a linear approximation around the fixed point, we found that it even provides a good estimate of the frequency of the sustained oscillation after the Hopf bifurcation.

We note that the parameters a, b have a large influence on the frequency. Figure 6-3 shows the frequency as a function of a and b. Here the parameters k_{ie} and k_{ei} are chosen such that the system is at the Hopf bifurcation point. This result means that by choosing different combinations of rate constants one can achieve a large range of frequency variation.

6.2.2 Two Coupled Columns

To investigate the dynamics of coupled columns we use the same parameters a, b, Q_m, k_{ie} and k_{ei}, for each column. The heterogeneity is introduced in the coupling gains between the excitatory populations of different columns. We start with the system given in Equation (6.5), with N = 2. The bifurcation scenarios seen in this case will be then generalized to the N-coupled case.

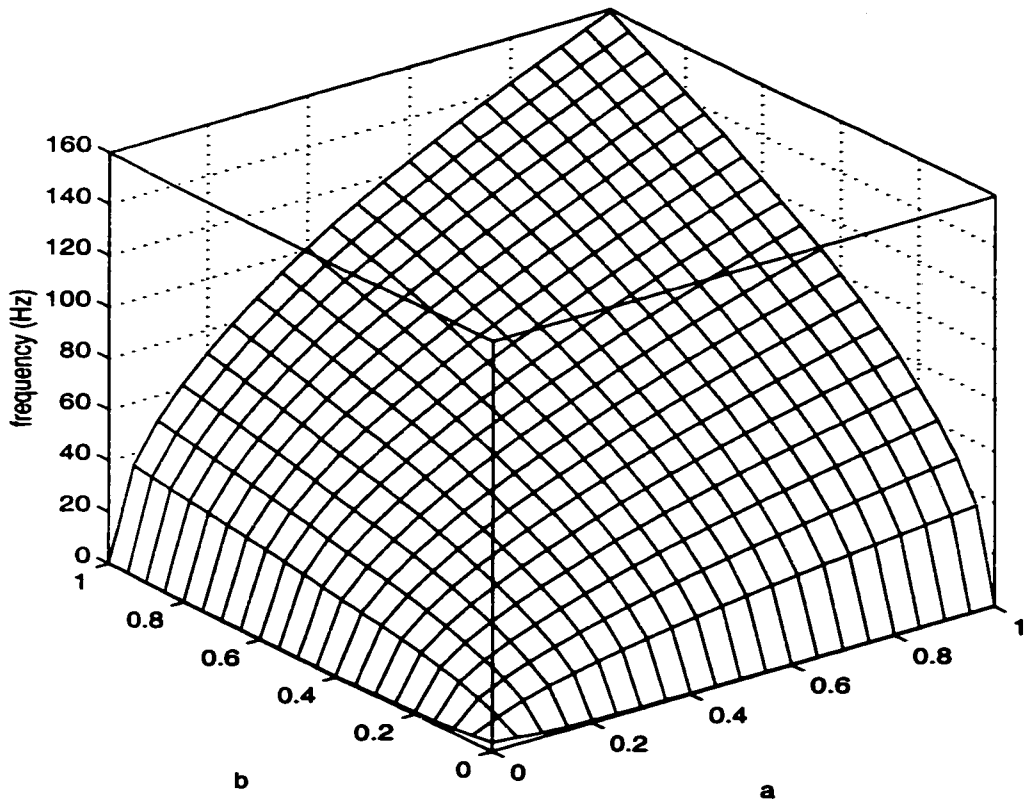

Figure 6-3: Frequency dependence on the rate constants a and b, obtained from the linear estimation for the frequency of the excitatory and inhibitory populations in the single column. The frequency estimation is computed from Equation (6.90), with the term $k_{ei}k_{ie} = ab(a+b)^2$. Rate constants given in milliseconds.

6.2.2.1 Properties of the Origin in the Absence of Input

Here we consider the dynamical properties of the fixed point in preparation for understanding its bifurcation in the presence of increasing input. We consider the case of $I_1 = I_2 = 0$. For simplicity we let $k_{21} = k_{12} = k$. The stability of the fixed point at the origin was investigated in previous section. Here, using a different approach by taking the

sum and difference of two column variables, we obtain the following characteristic equations for the linearized equations:

$$\begin{aligned}(\lambda+a)^2(\lambda+b)^2-k[\lambda^2+(a+b)\lambda+ab]+k_{ie}k_{ei}=0\\(\lambda+a)^2(\lambda+b)^2+k[\lambda^2+(a+b)\lambda+ab]+k_{ie}k_{ie}=0\end{aligned}. \tag{6.91}$$

Within the stability constraint given in the previous section, the presence of complex roots will indicate the oscillatory nature of the fixed point. This can then be viewed as a precursor to the Hopf bifurcation as input increases. Without attempting to derive formally the conditions for imaginary eigenvalues, to facilitate the analysis of the roots, we rewrite the above equations as

$$\begin{aligned}f(\lambda)-g(\lambda)=0\\f(\lambda)+g(\lambda)=0\end{aligned}, \tag{6.92}$$

where

$$\begin{aligned}f(\lambda)=(\lambda+a)^2(\lambda+b)^2+k_{ie}k_{ei}\\g(\lambda)=k[\lambda^2+(a+b)\lambda+ab]\end{aligned}. \tag{6.93}$$

Plotted on the same graph the intersections of the function f with the functions g and -g give the real roots of the characteristic equations. The number of real roots is significantly affected by the levels of interactions within the column and between columns. We mention that a particularly interesting scenario occur as follows. For a given set of values a, b, $k_{ie}k_{ei}$, when $k = 0$, the origin is oscillatory based on discussions on the singe column case. As k increases one may arrive at the situation depicted in Figure 6-4 where one finds 8 real roots, meaning that the oscillation around the fixed point is abolished by excitatory coupling between columns. As k further increases, the number of real roots becomes

smaller than 8 again, restoring the oscillatory properties at (0,0). For the purpose of this paper it suffices to say that high enough levels of interactions among constituent populations in a coupled column lead to oscillations around the fixed point at the origin.

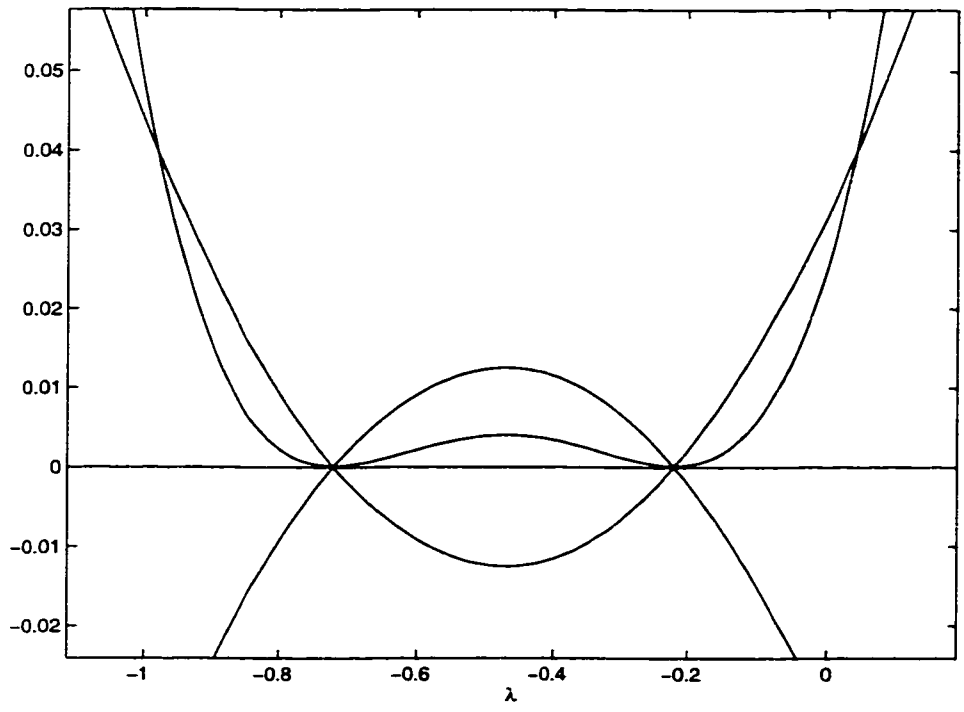

Figure 6-4: Functions f and g used for determining the nature of the equilibrium point at $(x, y) = (0,0)$ for the 2-symmetrically coupled columns. The intersections between the functions $f(\lambda)$, $g(\lambda)$ and $-g(\lambda)$ represent the real roots of the eigenvalue equation for the 2-coupled columns system. This figure is designed to demonstrate that for two coupled columns, unlike a single column system, the origin may not always be oscillatory. Specifically, for the situation depicted in the figure, all the 8 possible eigenvalues are real, meaning that the origin is non-oscillatory. But we comment that this is an extreme case where the interaction between the excitatory and inhibitory populations is set to a very low level. Increasing the coupling strengths k_{ie} and k_{ei}, or k will easily lead to complex roots, giving the fixed point an oscillatory response to perturbations. The existence of complex eigenvalues is a condition for the Hopf bifurcation. The parameters were set such that the fixed point is stable, with $a = 220/s, b = 720/s, k_{ie}k_{ei} = 0.001$ and $k = 0.1$.

6.2.2.2 Bifurcations Under Input

The homogeneous case: $k_{21} = k_{12}$ and $I_1 = I_2$.

Consider the synchronized solution. The system can be reduced to a 4th order system:

$$\begin{aligned} \frac{d^2x}{dt^2} + (a+b)\frac{dx}{dt} + abx &= -k_{ei}Q(y, Q_m) + kQ(x, Q_m) + I \\ \frac{d^2y}{dt^2} + (a+b)\frac{dy}{dt} + aby &= k_{ie}Q(x, Q_m) \end{aligned} \quad . \tag{6.94}$$

Let the fixed point be denoted by $x_1^* = x_2^* = x_0$ and $y_1^* = y_2^* = y_0$. The equation for the fixed point can be written as

$$\begin{aligned} x_0 &= \frac{1}{ab}\left(kQ(x_0, Q_m) - k_{ei}Q\left(\frac{k_{ie}Q(x_0, Q_m)}{ab}, Q_m\right) + I\right) \\ y_0 &= \frac{k_{ie}Q(x_0, Q_m)}{ab} \end{aligned} \quad . \tag{6.95}$$

It can be seen that x_0 is the root of the following function

$$\mathrm{f}(x) = abx + k_{ei}Q\left(\frac{k_{ie}Q(x, Q_m)}{ab}, Q_m\right) - kQ(x, Q_m) - I. \tag{6.96}$$

We plot f(x) as a function of x in Figure 6-5 for three sets of values of a, b, Q_m, k_{ie}, k_{ei} and $I = 0$. We note that, differently from the single column case, f(x) can be a non-monotonic function. This allows for multi-stability, saddle node bifurcations and hysteresis.

Based on the behavior of f(x) as a function of input I one can imagine the following scenarios regarding the fixed points in the system. (It should be noted that, concurrent to the changes in fixed point structures, one observes the bifurcations creating

or annihilating limit cycles.) Specifically, in Figure 6-5, curve (a), only one fixed point exists. As input increases the function f(x) is moved downward in a rigid fashion and this fixed point is moved to a higher value. Curve (b) implies that, as I increases, one observes the creation and destruction of fixed points, through saddle node bifurcations. At first, a saddle node bifurcation generates two additional fixed points, one stable and one unstable. As input increases further a reverse saddle node bifurcation happens, annihilating the original stable fixed point and the newly created unstable one. When this happens the system jumps to the only remaining fixed point at a higher excited state. This scenario allows for hysteresis. In curve (c), three fixed points coexist in the absence of input: a stable fixed point at the origin (low equilibrium point), an unstable one at an intermediate level and another stable one at a higher value (high equilibrium point). When input increases, a reverse saddle node bifurcation occurs through the annihilation of a stable and unstable fixed point.

The above scenarios concerning the fixed point structures provide only one factor to the understanding of the system's dynamics under input. Another factor is that Hopf bifurcations can take the dynamics away from the fixed point and this oscillatory dynamics can undergo its own set of bifurcations under input. In common bifurcation scenario, related to the curve (a) in Figure 6-5, as the input reaches higher and higher values, the fixed point of the system enters low gain (slope) regions of the sigmoid output function, and becomes stable again. A reverse Hopf bifurcation takes place. The amplitude of the oscillations shrinks until it terminates. Another more interesting cenario, shown in Figure 6-6, where the oscillatory behavior is terminated by a homoclinic

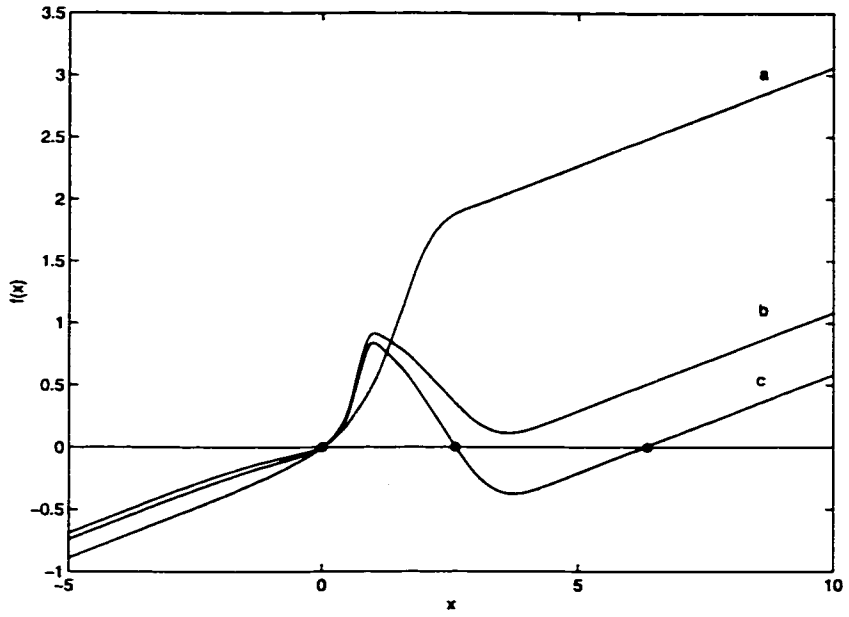

Figure 6-5: The fixed point equation for the 2-symmetrically coupled columns with input $I = 0$; The ploted variable is the local field potential of the excitatory population (x). Three main situations are plotted: in curve (a) the system has only one fixed point (represented by the intersection with the zero line). As the input level increases above zero, the curve moves downward (not shown) and the fixed point moves to a higher value of x (parameters: $Q_m = 5$, $k_{ie} = 0.1$, $k_{ei} = 0.4$, $k_{21} = k_{12} = 0.1$); In curve (b), besides the fixed point at the origin, the system has two new possible fixed points. As input increases, a saddle node bifurcation gives origin to one unstable and one stable fixed points (parameters: $Q_m = 10$, $k_{ie} = 0.4$, $k_{ei} = 0.1$, $k_{21} = k_{12} = 0.15$); In curve (c), three new fixed points are already in place for $I = 0$; The system has two stable equilibrium points, one at the origin and one at a higher level. In an intermediate level there an unstable fixed point (parameters: $Q_m = 10$, $k_{ie} = 0.4$, $k_{ei} = 0.1$, $k_{21} = k_{12} = 0.2$); See text for the relations between this possible attractor layouts and the Hopf bifurcation. In all the three previous cases $a = 220/s$, $b = 720/s$.

bifurcation (Strogatz, 1994). At the bifurcation point one should observe infinite long period oscillations. In our numerical simulations, the oscillations indeed become slower and slower. However, they terminate abruptly with the system jumping to a stable fixed

point. Our hypothesis is that the stable limit cycle approaches the invariant manifolds of the saddle. If the stable manifold is also the border of the basin of attraction separating the stable limit cycle and a stable fixed point, any small perturbation (numerical or physical) can make the system to enter the basin of attraction of the stable fixed point. This hypothesis requires then multistability, which is possible for $N \geq 2$. In Figure 6-6, the initial conditions for the system are at the origin. For $I = 0$, the 2-coupled columns system has already three fixed points. A stable one at the origin (low equilibrium state), another stable near $x_1 = 5$ (high equilibrium state) and an unstable one at an intermediary level (represented by the dotted line). The low equilibrium state becomes unstable at a certain level of input and a limit cycle is created. Here the bifurcation into two branches represents the maximum and minimum amplitude of the cycle. The oscillations are abruptly terminated in the form described. Note that when the input level is decreased from a higher level, the network dynamics follows a different route without the presence of oscillations. The multistability scenario can lead also to histeresis, a phenomenon typical of nonlinear systems having multiple coexisting attractors (multistability). One interesting implication of this scenario shown is that, although the system can start at the low equilibrium state, as the input follows its rise-and-fall temporal pattern the system might relax back to a higher equilibrium or excited state, and remain there in the absence of external input. The same type of scenario was also common for higher dimensional systems, e.g., a 64-coupled columns system.

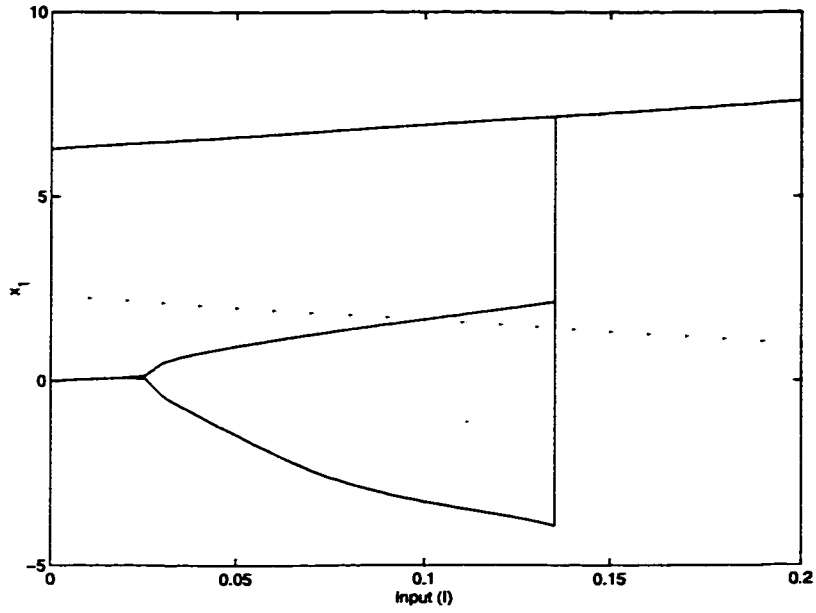

Figure 6-6: Bifurcation diagram for the 2 symmetrically coupled columns. The bifurcation parameter is the input level I and only the activity of the excitatory population x_1 is shown. Initial conditions for he system are at the origin. For $I = 0$ the 2 coupled column system already have three fixed point. A stable one at the origin (low equilibrium state), another stable near to $x_1 = 6$ (high equilibrium state) and an unstable one at intermediary level (represented by the dashed line). The low equilibrium state becomes unstable for a certain level of input a limit cycle cycle is created (bifurcation into upper and lower branches). The limit cycle is terminated not by a reverse Hopf, but by a homoclinic bifucation. At that point the 2-coupled column system jumps to the higher equilibrium state. By varying the input in the reverse direction (from high to zero) no oscillations are possible. The system stays all the way in the higher equilibrium state. This scenario relates to the curve (b) shown in Figure 6-5. The same parameters used there are applied here.

The heterogeneous case: $k_{21} \neq k_{12}$ and/or $I_1 \neq I_2$.

For the asymmetrically coupled model the same scenarios shown above are still present. Additionally, we observe quasiperiodicity and weak chaos over a narrow range of input that is not seen in the symmetric case (but we could not rule out their possible occurrence in the homogeneous case). The bifurcation diagram (Figure 6-7 and the

enlargement in Figure 6-8) shows three main regions: stable fixed point, periodic oscillations and a small intermediate interval where quasiperiodicity and chaos appear. The new dynamics are characterized by fast oscillations around 50 Hz with a slow amplitude modulation around 4Hz. Figure 6-9 presents an example of this oscillatory behavior in the chaotic regime. The computed correlation dimension for the time series in this regime is about 2.25 suggesting an attractor structure more complex than a torus (Strogatz, 1994). Careful examination of the bifurcation diagram also reveals periodic windows in which periodic solutions with more complex waveforms occur.

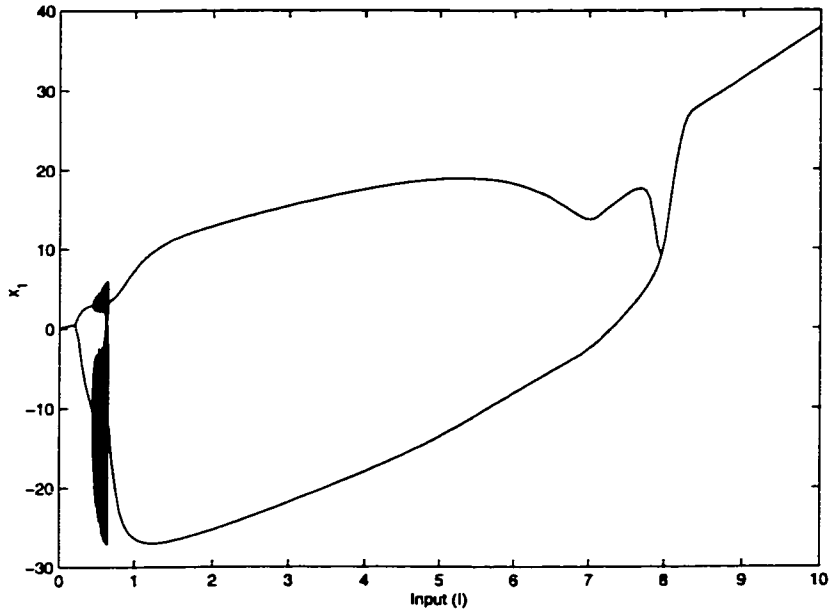

Figure 6-7: Bifurcation diagram for the 2 asymmetrically coupled columns. The bifurcation parameter is the input level I and only the activity of one the excitatory populations x_1 is shown. Besides the bifurcation into the limit cycle, a small region of periodic, quasiperiodic and chaotic behavior exists represented by the darkened region (see Figure 6-8 for the enlarged plot of this region). (Parameters: $a = 220/s$, $b = 720/s$, $Q_m = 20$, $k_{ie} = 0.1$, $k_{ei} = 0.4$, $k_{21} = 0.02$, $k_{12} = 0.2$);

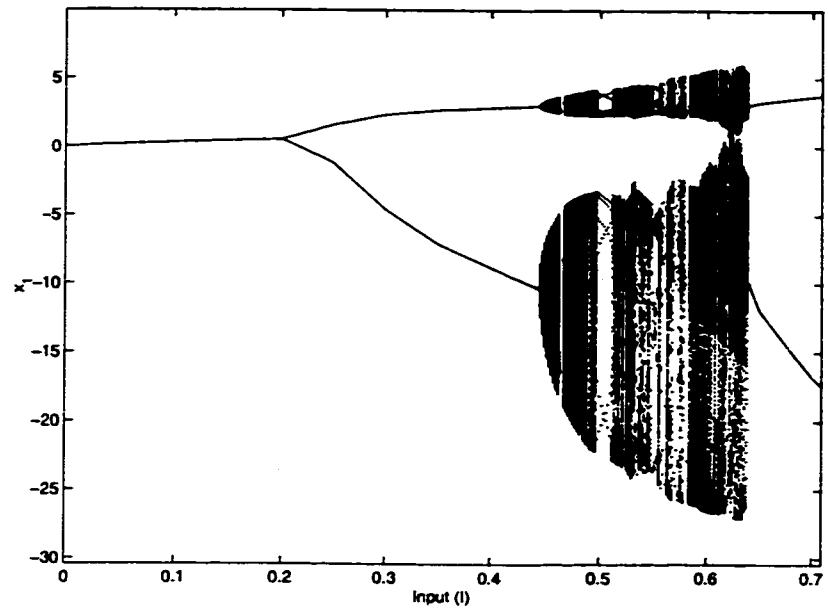

Figure 6-8: Bifurcation diagram for the 2 asymmetrically coupled columns. The bifurcation parameter is the input level I and only the activity of the excitatory population x_1 is shown. This plot shows an enlarged part of the previous Figure 6-7. In the region of input level from ~ 0.2 to ~ 4.5, a limit cycle exists. From I near to 4.5 to ~ 6.4, subregions of periodic, quasiperiodic and chaotic behavior coexist.

6.2.3 *N*-Coupled Columns

The system of N-coupled columns with homogeneous coupling $k_{np} = k$, $\forall n, p$ exhibits the same range of qualitative behavior as that for the 2-coupled columns with equal interconnections. Below we study a heterogeneous case where a coupling matrix consists of a random set of Gaussian variants.

6.2.3.1 Bifurcations Under Input

We consider a 64 coupled column model. The activity of only one of the excitatory populations is shown. The couplings k are randomly chosen according to a

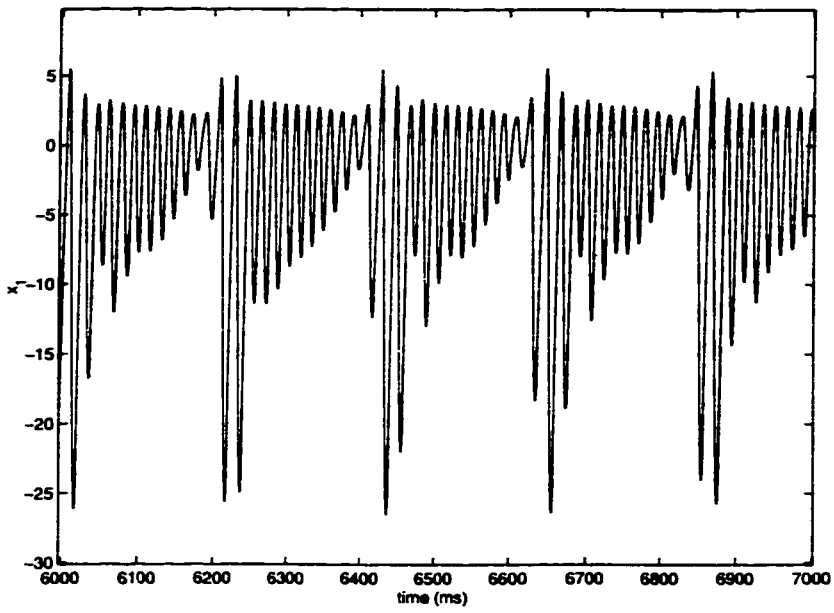

Figure 6-9: Chaotic behavior of the 2 asymmetrically coupled system. The activity of one of the excitatory populations x_1 is shown. Correlation dimension computed from a larger segment than the one shown presented here allowed values about 2.25 suggesting an attractor structure more complex than a torus or quasiperiodic behavior, pointing to a chaotic type of behavior. (Parameters: $a = 220/s$, $b = 720/s$, $Q_m = 20$, $k_{ie} = 0.1$, $k_{ei} = 0.4$, $k_{21} = 0.02$, $k_{12} = 0.2$, $I_1 = I_2 = 0.6$);

Gaussian distribution ($\mu = 0.15, \sigma = 0.025$). For small Q_m, the bifurcation diagram presents a stable fixed point and Hopf bifurcations. The 3 plots in Figure 6-10 are for the same set of couplings but Q_m is set at a higher value. Note the 3 distinct oscillatory behaviors: periodic (~40Hz), aperiodic (~ 50 Hz with a ~3Hz, input level I = 2.25), and periodic (~30Hz). The computation of the correlation dimension of the aperiodic time series gave values higher then 2, suggesting as in the previously shown case of 2 asymmetrically coupled columns, a structure more complex than a quasiperiodic attractor. The amplitude of the population's activity in each column is dependent on the total driving, i.e., external input and the input from the other columns. Because of the

heterogeneous coupling, the network presents a pattern of modulated local field potential amplitudes and phases over its columns (not shown).

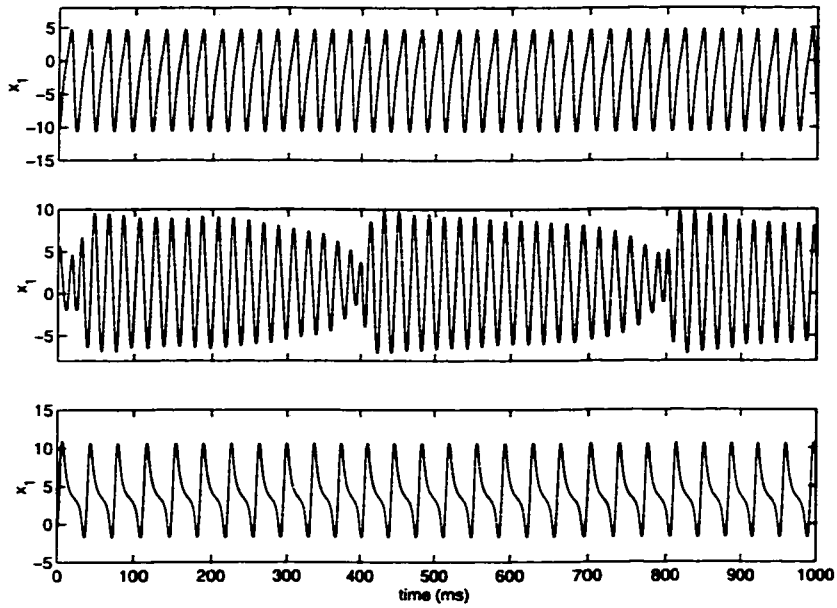

Figure 6-10: Types of behavior for the 64 asymmetrically coupled columns system. The coupling strength for the mutual excitation between the columns were taken from a Gaussian distribution ($\mu = 0.15, \sigma = 0.025$) and were randomly assigned to each connection. The plots show the activity of one of the excitatory populations x_1. Differently from the previous Figure 6-10, here Q_m is set to 10. The top plot shows a time series for input level $I = 0.5$ where the behavior is periodic (~ 40 Hz). The middle plot shows the activity for $I = 2.25$. Correlation dimension computed from larger segments than the one shown here allowed values higher than 2, suggesting an attractor structure more complex than a torus, i.e., chaotic behavior. The oscillations around 50 Hz with a slower modulation around 3 Hz resembles a beating phenomenon. This region of nonperiodic behavior exists for a larger interval of input levels than the one for the 2 asymmetrically coupled system. This region of aperiodic behavior has coexistent subregions of quasiperiodic and chaotic activity. The system becomes periodic (~ 30 Hz) again for a higher level of input (bottom plot, $I = 3.1$); Further increase of the input level leads to the destruction of the limit cycle through a homoclinic bifurcation. The oscillations suddenly stop, instead of gradually decreasing to zero as in a reverse Hopf bifurcation.

References

Aertsen, A., Erb, M., & Palm, G. (1994). Dynamics of functional coupling in the cerebral cortex: an attempt at a model-based interpretation. *Physica D, 75*, 103-128.

Aertsen, A., Gerstein, G.L., Habib, M.K., & Palm, G. (1989). Dynamics of neuronal firing correlation: modulation of "effective connectivity". *Journal of Neurophysiology, 61*(5), 900-917.

Arieli, A., Sterkin, A., Grinvald, A., & Aertsen, A. (1996). Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. *Science, 273*, 1868-1871.

Astrom, K.J. (1970). *Introduction to stochastic control theory*. New York: Academic Press.

Atiya, A.F. (1988). Learning on a general network. In D.A. Anderson (Ed.), *Neural Information Processing Systems* . New York: American Institute of Physics.

Attick, J.J. (1992). Could information-theory provide an ecological theory of sensory processing. *Network: Computation in Neural Systems, 3*(2), 213-251.

Azouz, R., & Gray, C.M. (1999). Cellular mechanisms contributing to response variability of cortical neurons in vivo. *The Journal of Neuroscience, 19*(6), 2209-2223.

Baccala, L.A., & Sameshima, K. (2001). Partial directed coherence: a new concept in neural structure determination. *Biological Cybernetics, 84*(6), 463-474.

Baird, B., & Eeckman, F.H. (1993). A Hierarchical Sensory-Motor Architecture of Oscillating Cortical Area Subnetworks. In F. Eeckman (Ed.), *Analysis and Modeling of Neural Systems* (pp. 96-104). Norwell, MA: Kluwer.

Basar, E., Gonder, A., & Ungan, P. (1976). Important relation between EEG and brain evoked potentials. I. Resonance phenomena in subdural structures of the cat brain. *Biological Cybernetics, 25*, 27-40.

Basar, E., Rahn, E., Demiralp, T., & Schurmann, M. (1998). Spontaneous EEG theta activity controls frontal visual evoked potentials. *Electroencephalography and Clinical Neurophysiology, 108*, 101-109.

Bell, A.J., & Sejnowski, T.J. (1995). An information maximization approach to blind separation and blind deconvolution. *Neural Computation, 7*(6), 1129-1159.

Beran, J. (1994). *Statistics for long-memory processes*. New York: Chapman & Hall.

Bienenstock, E., Geman, S., & Potter, D. (1997). *Composiitonality, MDL priors and object recognition*. Advances in Neural Information Processing Systems (NIPS).

Box, G.E.P., Jenkins, G.M., & Reinsel, G.C. (1994). *Time series analysis: Forecasting and control*. (3 ed.). New Jersey: Prentice Hall.

Brandt, M.E., & Jansen, B.H. (1991). The relationship between prestimulus alpha amplitude and visual evoked potential amplitude. *International Journal of Neuroscience, 61*, 261-268.

Bressler, S.L. (1995). Large-scale cortical networks and cognition. *Brain Research Rev., 20*, 288-304.

Bressler, S.L. (1996). Interareal synchronization in the visual cortex. *Behavioral Brain Research, 76*, 37-49.

Bressler, S.L., Coppola, R., & Nakamura, R. (1993). Episodic multiregional cortical coherence at multiple frequencies during visual task performance. *Nature, 366*, 153-156.

Bressler, S.L., Ding, M., & Liang, H. (1999). Investigation of cooperative cortical dynamics by multivariate autoregressive modeling of event-related local field potentials. *Neurocomputing, 26*(7), 625-631.

Brillinger, D.R. (1975). *Time series analysis: data analysis and theory*. New York: Holt, Rinehart & Winston.

Brody, C. (1998). Slow covariations in neuronal resting potentials can lead to artefactually fast cross-correlations in their spike trains. *Journal of Neurophysiology, 80*(6), 3345-3351.

Buchel, C., & Friston, K.J. (1997). Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI. *Cerebral Cortex, 7*, 768-778.

Burg (1975). *Maximum entropy spectral analysis.* Ph.D. Dissertation, Stanford University.

Chatfield, C. (1995). *The analysis of time series.* New York: Chapman & Hall.

Chawla, D., Lumer, E.D., & Friston, K.J. (1999). The relationship between synchronization among neuronal populations and their mean activity levels. *Neural Computation, 11*, 1389-1411.

Chawla, D., Lumer, E.D., & Friston, K.J. (2000). Relating macroscopic measures of brain activity to fast, dynamic neuronal interactions. *Neural Computation, 12*(12), 2805-2821.

Chawla, D., Rees, G., & Friston, K.J. (1999). The physiological basis of attentional modulation in extrastriate visual areas. *Nature Neuroscience, 2*(7), 671-676.

Chen, M.-H., Shao, Q.-M., & Ibrahim, J.G. (2000). *Monte Carlo methods in Bayesian computation.* New York: Springer Verlag.

Chen, T., & Amari, S. (2001). New theorems on global convergence of some dynamical systems. *Neural Networks, 14*, 251-255.

Cohen, M.A., & Grossberg, S. (1983). Absolute stability of global pattern information and parallel memory storage by competitive neural networks. *IEEE Transactions SMC, 13*(815-826).

Coppola, R., Tabor, R., & Buchsbaum, M.S. (1978). Signal to noise ratio and response variability measurements in single trial evoked potentials. *Electroencephalograpy and Clinical Neurophysiology, 44*, 214-222.

Cover, T.M., & Thomas, J.A. (1991). *Elements of information theory*. New York: John Wiley & Sons.

Damasio, A. (1989a). The brain binds entities and events by multiregional activation from convergence zones. *Neural Computation, 1*, 123-132.

Damasio, A.R. (1989b). Time-locked multiregional retroactivation: a systems-level proposal for the neural substrates of recall and recognition. *Cognition, 33*, 25-62.

Dawson, G.D. (1954). A summation technique for the detection of small evoked potentials. *Electroencephalography and Clinical Neurophysiology, 6*, 153-154.

Dayan, P., & Abbott, L.F. (2001). *Theoretical Neuroscience. Computational and mathematical modeling of neural systems*. Cambridge: MIT Press.

Deco, G., & Schurmann, B. (2000). *Information dynamics. Foundations and applications*. New York: Springer-Verlag.

Ding, M., Bressler, S.L., Yang, W., & Liang, H. (2000). Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: data preprocessing, model validation, and variability assessment. *Biological Cybernetics, 83*, 35-45.

Dong, D.W., & Atick, J.J. (1995). Temporal decorrelation - a theory of lagged and nonlagged responses in the lateral geniculate nucleus. *Network: Computation in Neural Systems, 6*(2).

Eeckman, F.H., & Freeman, W.J. (1991a). Asymmetric sigmoid nonlinearity in the rat olfactory system. *Brain Research, 557*, 13-21.

Eeckman, F.H., & Freeman, W.J. (1991b). Asymmetric Sigmoid Nonlinearity in the Rat Olfactory System. *Brain Research, 557*(1-2), 13-21.

Fang, Y., & Kincaid, T.G. (1996). Stability analysis of dynamical neural networks. *IEEE Transactions in Neural Networks, 7*, 996-1006.

Fisher, N.I. (1995). *Statistical analysis of circular data.* Cambridge: Cambridge University Press.

Forti, M., & Tsei, A. (1995). New conditions for global stability of neural networks with application to linear and quadratic programming problems. *IEEE Transactions in Circuit and Systems I, 42*, 354-366.

Franaszczuk, P.J., & Blinowska, K.J. (1985). Linear model of brain electrical activity - eeg as a superposition of damped oscillatory modes. *Biological Cybernetics, 53*, 19-25.

Freeman, W.J. (1975). *Mass action in the nervous system.* New York: Academic Press.

Freeman, W.J. (1979). Nonlinear gain mediating cortical stimulus-response relations. *Biological Cybernetics, 33*, 237-247.

Freeman, W.J. (1992). Tutorial on neurobiology: from single neurons to brain chaos. *International Journal of Bifurcation and Chaos, 2*(3), 451-482.

Freeman, W.J. (1994). Neural Mechanisms Underlying Destabilization of Cortex By Sensory Input. *Physica D, 75*(1-3), 151-164.

Fries, P., Reynolds, J.H., Rorie, A.E., & Desimone, R. (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. *Science, 291*(5508), 1560-1563.

Friston, K.J. (2000). The labile brain. I. Neuronal transients and nonlinear coupling. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, *355*(1394), 215-236.

Gantmacher, F.R. (1964). *Theory of matrices*. New York: Chelsea.

Granger, C.W.J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, *37*, 424-438.

Gray, C.M. (1999). The temporal correlation hypothesis of visual feature integration: still alive and well. *Neuron*, *24*(1), 31-47.

Gray, C.M., Konig, P., Engel, A.K., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit intercolumnar synchronization which reflects global stimulus properties. *Nature*, *338*, 334-337.

Guan, Z.H., Chen, G., & Qin, Y. (2000). On equilibria, stability, and instability of Hopfield neural networks. *IEEE Transactions in Neural Networks*, *11*, 534-540.

Guez, A., Protopopsecu, V., & Barhen, J. (1988). On the stability, storage capacity and design of nonlinear continuous neural networks. *IEEE Transactions in SMC*, *18*, 80-87.

Gur, M., Beylin, A., & Snodderly, D.M. (1997). Response variability of neurons in primary visual cortex (V1) of alert monkeys. *The Journal of Neuroscience*, *17*(8), 2914-2920.

Haykin, S., & Kesler, S. (1983). Prediction-Error Filtering and Maximum-Entropy Spectral Estimation. In S. Haykin (Ed.), *Topics in Applied Physics: Nonlinear Methods of Spectral Analysis* (Vol. 34, pp. 9-70). New York: Springer-Verlag.

Hinton, G.E., & Ghahramani, Z. (1997). Generative models for discovering sparse distributed representations. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 352*(1358), 1177-1190.

Hirsch, M.W. (1989). Convergent activation dynamics in continuous time network. *Neural Networks, 2*, 331-349.

Hopfield, J.J., & Brody, C.D. (2001). What is a moment? Transient synchrony as a collective mechanism for spatiotemporal integration. *Proceedings of the National Academy fo Sciences fo the United States of America, 98*(3), 1282-1287.

Hoppensteadt, F.C., & Izhikevich, E.M. (1999). Oscillatory neurocomputers with dynamic connectivity. *Physical Review Letters, 82*(14), 2983-2986.

Horn, R.A., & Johnson, C.R. (1990). *Matrix analysis*. New York: Cambridge University Press.

Horvath, R.S. (1969). Variability of cortical auditory evoked response. *Journal of Neurophysiology, 32*, 1056-1063.

Isard, M., & Blake, A. (1998). Condensation - conditional density propagation for visual tracking. *International Journal of Computer Vision, 29*(1), 5-28.

Izhikevich, E.M. (1999). Weakly connected quasi-periodic oscillators, FM interactions, and multiplexing in the brain. *Siam Journal On Applied Mathematics, 59*(6), 2193-2223.

Jaskowski, P., & Verleger, R. (1999). Amplitude and latencies of single-trial ERP's estimated by a maximum-likelihood method. *IEEE Transactions on Biomedical Engineering, 46*(8), 987-993.

Jaskowski, P., & Verleger, R. (2000). An evaluation of methods for single-trial estimation of P3 latency. *Psychophysiology, 37*(2), 153-162.

Jaynes, E.T. (1957b). Information theory and statistical mechanics. *Physical Review, 106*, 620-630.

Jaynes, E.T. (1957a). Information theory and statistical mechanics. *Physical Review, 108*(171-190).

Jaynes, E.T. (1994). Probability theory: the logic of science.

Jung, T.-P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., & Sejnowski, T.J. (1999). *Analyzing and visualizing single-trial event-related potentials.* Advances in Neural Information Processing Systems, NIPS.

Kalcher, J., & Pfurtscheller, G. (1995). Discrimination between phase-locked and non-phase-locked event-related EEG activity. *Electroencephalography and Clinical Neurophysiology, 94*, 381-384.

Kaminski, M., Ding, M., Truccolo, W.A., & Bressler, S.L. (2001). Evaluating causal relations in neural systems: Granger Causality, Directed Transfer Function (DTF) and statistical assessment of significance. *Biological Cybernetics, 85*, 145-157.

Kay, S.M. (1988). *Modern spectral estimation.* Englewood Cliffs: Prentice Hall.

Kelly, D.G. (1990). Stability in contractive nonlinear neural networks. *IEEE Transaction in Biomedical Engineering, 37*, 231-242.

Kelso, J.A.S. (1995). *Dynamic patterns: the self-organization of brain and behavior.* Cambridge: MIT Press.

Kisley, M.A., & Gerstein, G.L. (1999). Trial-to-trial variability and state-dependent modulation of auditory-evoked responses in cortex. *Journal of Neuroscience, 19*(23), 10451-10460.

Knuth, K.H. (1997). Difficulties applying recent blind source separation techniques to EEG and MEG. In G.J. Erickson, J.T. Rychert, & C.R. Smith (Eds.), *Maximum Entropy and Bayesian Methods* (pp. 209-222). Boise, Idaho, USA: Kluwer, Dordrecht, 1998.

Knuth, K.H. (1998). Bayesian source separation and localization. In A. Mohammad-Djafari (Ed.), *SPIE* (Vol. 3459, pp. 147-158).

Knuth, K.H. (1999). A Bayesian approach to source separation. *International Workshop on Independent Component Analysis* (pp. 283-288).

Knuth, K.H., Truccolo, W.A., Bressler, S.L., & Ding, M. (submitted). Separation of multiple evoked components using differential amplitude and latency variability. , *Third International Conference on Independent Component Analysis and Blind Signal Separation, 2001*. San Diego, California, USA.

Knuth, K.H., & Vaughan, H.G.J. (1999). Convergent Bayesian formulations of blind source separation and electromagnetic source estimation. In W. von der Linden, V. Dose, R. Fischer, & R. Preuss (Eds.), *Maximum Entropy and Bayesian Methods* (pp. 217-226). Garching, Germany: Kluwer Academic Publishers.

Lange, D.H., Pratt, H., & Inbar, G.F. (1997). Modeling and estimation of single evoked brain potential components. *IEEE Transactions on Biomedical Engineering, 44*(9), 791-799.

Li, Z., & Hopfield, J.J. (1989). Modeling the olfactory-bulb and its neural oscillatory processing. *Biological Cybernetics, 61*(5), 378-392.

Liang, X.B., & Wu, L.D. (1998). New sufficient conditions for absolute stability of neural networks. *IEEE Transactions in Circuits and Systems I, 45*, 584-586.

Mackay, D.J.C. (1992). The evidence framework applied to classification networks. *Neural Computation, 4*(5), 720-736.

Mangun, G.R., & Hillyard, S.A. (1991). Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual spatial priming. *Journal of Experimental Psychology - Human Perception and Performance, 17*(4), 1057-1074.

Matsuoka, K. (1991). On absolute stability of neural networks. *Transactions of the IECE (Japan), J74-D-11*, 536-542.

McGillem, C.D., & Aunon, J.I. (1987). Analysis of event-related potentials. In A.S. Gevins & A. Remond (Eds.), *Methods of analysis of brain electrical and magnetic signals. EEG Handbook* (Vol. 1, pp. Chapter 5): Elsevier Science Publishers.

Michel, A.N., Farrell, J.A., & Porod, W. (1989). Qualitative analysis of neural networks. *IEEE Transactions in Circuits and Systems, 36*, 229-243.

Mocks, J., Gasser, T., Tuan, P.D., & Kohler, W. (1987). Trial-to-trial variability of single potentials: methodological concepts and results. *International Journal of Neuroscience, 33*, 25-32.

Morf, M., Vieira, A., Lee, D., & Kailath, T. (1978). Recursive multichannel maximum entropy spectral estimation. *IEEE Transactions in Geoscience and Electronics, 16*, 85-94.

Nikias, C.L., & Petropulu, A.P. (1993). *Higher-order spectral analysis: a nonlinear signal processing framework.* New Jersey: Prentice and Hall.

Nowak, L.G., & Bullier, J. (1997). The timing of information transfer in the visual system. In K.S. Rockland (Ed.), *Cerebral Cortex* (Vol. 12, pp. 205-241). New York: Plenum Press.

Olshausen, B.A., Anderson, C.H., & Van Essen, D.C. (1993). A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *The Journal of Neuroscience, 13*(11), 4700-4719.

Olshausen, B.A., & Field, D.J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature, 381*(6583), 607-609.

Percival, D.B., & Walden, A.T. (1993). *Spectral Analysis for Physical Applications.* New York: Cambridge University Press.

Pfurtscheller, G., & Lopes da Silva, F.H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology, 110*, 1842-1857.

Pham, D.T., Mocks, J., Kohler, W., & Gasser, T. (1987). Variable latencies of noisy signals: estimation and testing in brain potential data. *Biometrika, 74*(3), 525-533.

Rangarajan, G., & Ding, M. (2000). Integrated approach to the assessment of long range correlation in time series data. *Physical Review E, 61*(5), 4991-5001.

Rissanen, J. (1989). *Stochastic complexity and statistical inquiry* (Vol. 15). New Jersey: World Scientific.

Rodriguez, E., George, N., Lachaux, J.-P., Martinerie, J., Renault, B., & Varela, F.J. (1999). Perception's shadow: long-distance synchronization of human brain activity. *Nature, 397*, 430-433.

Shah, A.S., Knuth, K.H., Mehta, A.D., Fu, K.G., Johnston, T.A., Dias, E.C., Truccolo, W.A., Ding, M., Bressler, S.L., & Schroeder, C.E. (2001). Functional connectivity between visual structures in behaving monkeys. *Society for Neuroscience Annual Meeting, San Diego, abstract.*

Sivia, D.S. (1996). *Data analysis: a Bayesian tutorial.* New York: Oxford University Press.

Sporns, O., Tononi, G., & Edelman, G.M. (1994). Reentry and dynamical interactions of cortical networks. In E. Domany, J.L. van Hemmen, & K. Schulten (Eds.), *Models of Neural Networks II* (pp. 315-341). New York: Springer-Verlag.

Srinivasan, R., Russell, D.P., Edelman, G.M., & Tononi, G. (1999). Increased synchronization of neuromagnetic responses during conscious perception. *Journal of Neuroscience, 19*(13), 5435-5448.

Strogatz, S.H. (1994). *Nonlinear dynamics and chaos. With applications to physics, biology, chemistry and engineering.* Reading: Addison-Wesley.

Sugawara, K., Harao, M., & Noguchi, S. (1983). On the stability of equilibrium states of analogue neural networks. *Transactions of the IECE (Japan), J66-A*, 258-265.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1997). Oscillatory gamma-band (30-70 Hz) activity induced by a visual search task in humans. *The Journal of Neuroscience, 17*(2), 722-734.

Tallon-Baudry, C., Bertrand, O., Peronnet, F., & Pernier, J. (1998). Induced gamma-band activity during the delay of a visual short-term memory task in humans. *Journal of Neuroscience, 18*(11), 4244-4254.

Thomson, D.J. (1982). Spectrum estimation and harmonic analysis. *Proceedings of the IEEE, 70*, 1055-1096.

Thomson, D.J. (2000). Multitaper analysis of nonstationary and nonlinear time series. In W. Fitzgerald, R. Smith, A. Walden, & P. Young (Eds.), *Nonlinear and nonstationary signal processing* . New York: Cambridge University Press.

Tononi, G., Sporns, O., & Edelman, G.M. (1992). Reentry and the problem of integrating multiple cortical areas: Simulation of dynamic integration in the visual system. *Cerebral Cortex, 2*(4), 310-335.

Truccolo, W.A., & Dong, D.W. (2001). Dynamic temporal decorrelation: an information-theoretic and biophysical model the functional role of the lateral geniculate nucleus. *Neurocomputing, 38-40*, 993-1001.

Truccolo, W.A., Knuth, K.H., Bressler, S.L., & Ding, M. (in press). Bayesian analysis of single trial cortical evoked components. In R. Fry & M. Bierbaum (Eds.),

Maximum Entropy and Bayesian Methods . Johns Hopkins University, Baltimore: Kluwer.

Vijn, P.C.M., van Dijk, B.W., & Spekreijse, H. (1991). Visual-stimulation reduces EEG activity in man. *Brain Research, 550*(1), 49-53.

von der Malsburg, C. (1994). The correlation theory of brain function. In E. Domany, J.L. van Hemmen, & K. Schulten (Eds.), *Models of Neural Networks II* (pp. 95-119). New York: Springer-Verlag.

Wilson, H., & Cowan, J.D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Byophysical Journal, 12*, 1-24.

Wold, H. (1938). *A study in the analysis of stationary time series.* Stockholm: Almqvist & Wiksell.

Woody, C.D. (1967). Characterization of an adaptive filter for the analysis of variable latency neuroelectric signals. *Medical Biological Engineering, 5*, 539-553.

Yang, H., & Dillon, T.S. (1994). Exponential stability and oscillation of Hopfield graded response neural network. *IEEE Transactions in Neural Networks, 5*, 719-729.

Zemel, R.S. (1994). *A minimum description length framework for unsupervised learning,* Ph.D. Dissertation, University of Toronto.

Curriculum Vitae

NAME: Wilson A. Truccolo-Filho

CITIZENSHIP: Dual Brazil/Italy

ACADEMIC BACKGROUND .....

Ph.D. in Complex Systems and Brain Sciences (1996, expected conclusion: 2001), Advisors: Drs. Steven Bressler and Mingzhou Ding, Center for Complex Systems and Brain Sciences, Florida Atlantic University.

M.Sc. in Cognitive Psychology (1994-1995): Oscillatory Behavior of Neuronal Populations and Cognitive Processes: an Electroencephalographic Study. Department of Psychology, Federal University of R. Grande do Sul (UFRGS), and Department of Biomedical Engineering, Clinical Hospital of Porto Alegre, Brazil.

Undergraduate Studies in Mathematics (1994-1995), Department of Mathematics, (UFRGS). Courses: Calculus, Linear Algebra, Differential Equations, Real Analysis.

Graduate studies in Philosophy (1990). Phenomenology and Hermeneutics, Heidegger and Wittgenstein, with Professor Dr. Ernildo Stein and Luis C. Petry, Department of Philosophy, UFRGS.

B.A. in Psychology: UFRGS, 1989.

COURSES .....

Summer Workshop on the Analysis of Neural Data, Marine Biology Laboratory (MBL), Woods Hole, MA, August 2001. Organizers: Partha Mitra (Lucent Technologies – Bell Laboratories), David Kleinfeld (University of California, San Diego) and Emery Brown (Massachusetts General Hospital).

FELLOWSHIPS AND AWARDS .....

Fellowship from the Office of Naval Research (ONR), 2000 – current.
Fellowship from the Brazilian Research Council CNPq, 1996-2000.
Newell Doctoral Fellowship, Florida Atlantic University, 2000.
Newell Doctoral Fellowship, Florida Atlantic University, 1999.
Fellowship from the Brazilian Research Council CNPq, 1994-1995.

TEACHING EXPERIENCE .....

Teaching Assistant (Spring, 2001): Center for Complex System and Brain Sciences, graduate courses:

Multivariate Signal Processing (Professor: Dr. Mingzhou Ding) and

Computational Neuroscience (Professor: Dr. Dawei Dong).

Teaching Assistant (Fall, 1995): Cognitive Psychology, Department of Psychology, UFRGS, Brazil, undergraduate course.

Teacher (1990): Peircean Semiotics, jesuitic University Vale do Rio dos Sinos (UNISINOS), Brazil, undergraduate course.

PUBLICATIONS .....

ANALYSIS OF NEURAL DATA

Truccolo, W.A., Ding, M., Knuth, K.H., Bressler, S.L. *(submitted)*. Trial-to-Trial Variability of Cortical Evoked Responses: Implications for the Analysis of Functional Connectivity.

Truccolo, W.A., Knuth, K.H., Bressler, S.L., Ding, M. (*in press*) Bayesian Analysis of Single Trial Cortical Evoked Components. In: Robert L. Fry and Michelle Bierbaum (Eds.), Maximum Entropy and Bayesian Methods: Johns Hopkins University, Baltimore, Kluwer.

Knuth, K.H., Truccolo, W.A., Bressler, S.L., Ding, M. (*in press*) Separation of multiple evoked responses using differential amplitude and latency variability. In: Sejnowski, T.J. (Ed.), Proceedings of the 3rd International Conference on Independent Component Analysis and Blind Signal Separation, December, 2001, San Diego, California, USA.

Kaminski, M., Ding, M., Truccolo, W.A., Bressler, S.L. (2001). Evaluating Causal Relations in Neural Systems: Granger Causality, Direct Transfer Function (DTF) and Statistical Assessment of Significance. *Biological Cybernetics*, 85, pp. 145-157.

Truccolo, W.A., Ding, M., Bressler, S.L. (2001). Variability and Interdependence of Local Field Potentials: Effects of Gain Modulation and Nonstationarity. *Neurocomputing*, 38-40, pp. 983-992.

COMPUTATIONAL NEUROSCIENCE AND NEURAL DYNAMICS

Truccolo, W.A., Dong, D.W. (2001). Dynamic Temporal Decorrelation: Information Theoretic and Biophysical Model of the Functional Role of the Lateral Geniculate Nucleus (LGN). *Neurocomputing*, 38-40, pp. 993-1001.

Truccolo, W.A., Ding, M., Bressler, S.L. (2000). Stability Constraints for Oscillatory Neural Networks. *Neurocomputing*, 32-33, pp. 585-589.

Truccolo, W.A., Rangarajan, G., Ding, M. *(submitted)* Coupling Strength Constraints for Equilibrium Points in Oscillatory Neural Networks.

Truccolo, W.A., Ding, M., Bressler, S.L. (2000). Dynamical Analysis of an Oscillatory Cortical Neural Network. In: Amari et al (Eds). Proceedings of the IEEE-INNS-ENNS-ENNS International Joint Conference on Neural Networks, Como, Italy, 2000.

Truccolo, W., Dong, D.W. (2000). Dynamic temporal decorrelation: Information theoretic and biophysical model of the functional role of lateral geniculate nucleus (LGN). *Investigative Ophthalmology & Visual Science*, 41 (4): 259B259, Suppl. S MAR 15 2000.

CONFERENCES.....

Truccolo, W.A., Knuth, K.H., Ding, M., Bressler, S.L. (2001). Bayesian Estimation of Amplitude, Latency and Waveform of Single Trial Cortical Evoked Components. INTERNATIONAL WORKSHOP ON BAYESIAN INFERENCE AND MAXIMUM ENTROPY METHODS IN SCIENCE AND ENGINEERING (MaxEnt, August 4-9, 2001), Johns Hopkins University, Baltimore, MD.

Knuth, K.H., Truccolo, W.A., Bressler, S.L., Ding, M. (2001). Separation of Multiple Evoked Responses Using Differential Amplitude and Latency Variability. 3rd International Conference on INDEPENDENT COMPONENT ANALYSIS and BLIND SIGNAL SEPARATION, December 9-13, 2001, San Diego, California, USA.

Truccolo, W.A., Knuth, K.H., Ding, M., Bressler, S.L. (2001). Bayesian Estimation of Single Trial Cortical Evoked Potential Components: Applications to the Analysis of Functional Connectivity. SOCIETY FOR NEUROSCIENCE ANNUAL MEETING, San Diego.

Albo, Z., Viana Di Prisco, G., Truccolo, W.A., Vertes, R.P., Ding. M. (2001). A Study of Neural Interactions Within the Limbic System using Partial Coherence and Direct Transfer Functions Analysis. SOCIETY FOR NEUROSCIENCE ANNUAL MEETING, San Diego.

Knuth, K.H., Truccolo, W.A., Shah, A.S., Ding, M., Bressler, S.L., Schroeder, C.E. (2001). Facing Up to Trial-to-Trial Variability of Evoked Responses. SOCIETY FOR NEUROSCIENCE ANNUAL MEETING, San Diego.

Shah, A.S., Knuth, K.H., Mehta, A.D., Fu, K.G., Johnston, T.A., Dias, E.C., Truccolo, W.A., Ding, M., Bressler, S.L., Schroeder, C.E. (2001). Functional Connectivity between Visual Structures in Behaving Monkeys. SOCIETY FOR NEUROSCIENCE ANNUAL MEETING, San Diego.

Truccolo, W.A., Ding, M., Bressler, S.L. (2000). Variability of Cortical Evoked Responses: Implications for the Analysis of Functional Connectivity. SOCIETY FOR NEUROSCIENCE ANNUAL MEETING, New Orleans.

Truccolo, W.A., Ding, M., Bressler, S.L. (2000). Dynamical Analysis of an Oscillatory Cortical Neural Network. INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS (IJCNN), Como, Italy.

Truccolo, W.A., Ding, M., Bressler, S.L. (2000). On the Variability of Cortical Evoked Responses: Implications for the Correlation and Spectral Coherence Analysis of Local Field Potentials. COMPUTATIONAL NEUROSCIENCE MEETING, Brugge, Belgium.

Truccolo, W.A. and Dong, D.W. (2000). Dynamic Temporal Decorrelation: Information Theoretic and Biophysical Model of the Functional Role of the Lateral Geniculate Nucleus (LGN). COMPUTATIONAL NEUROSCIENCE MEETING, Brugge, Belgium.

Truccolo, W.A., Ding, M., Bressler, S.L. (2000). Dynamics of Oscillatory Neural Networks. WORLD CONFERENCE ON SYSTEMICS, CYBERNETICS, AND INFORMATICS, Orlando.

Truccolo, W.A. and Dong, D.W. (2000). Dynamic Temporal Decorrelation: Information Theoretic and Biophysical Model of the Functional Role of the Lateral Geniculate Nucleus (LGN). ARVO (THE ASSOCIATION FOR RESEARCH IN VISION AND OPHTHAMOLOGY, ANNUAL MEETING).

Truccolo, W.A., Ding, M., Bressler, S.L. (1999). Stability of Oscillatory Neural Networks. DYNAMIC NEUROSCIENCE VII, INTEGRATION ACROSS MULTIPLE IMAGING MODALITIES, NEUROSCIENCE MEETING SATELLITE SYMPOSIUM, NIHM, Boca Raton.

Truccolo, W.A., Ding, M., Bressler, S.L. (1999). Variability of Cortical Evoked Responses: Implications for Cross-Correlation and Coherence Measures. DYNAMIC NEUROSCIENCE VII, INTEGRATION ACROSS MULTIPLE IMAGING MODALITIES, NEUROSCIENCE MEETING SATELLITE SYMPOSIUM, NIHM, Boca Raton.

Truccolo, W.A., Ding, M., Bressler, S.L. (1999). Dynamical Analysis of Oscillatory Neural Networks. 5th EXPERIMENTAL CHAOS CONFERENCE, Orlando, Florida.

Truccolo, W.A., Ding, M., Bressler, S.L. (1999). Dynamical Analysis of a Generic Cortical Area Model. COMPUTATIONAL NEUROSCIENCE MEETING, Pittsburg.

Truccolo, W.A., Ding, M., Bressler, S.L. (1998). Information Processing in a Generic Cortical Area Model. NEW ENGLAND COMPLEX SYSTEM CONFERENCE, Nashua, N.H.

PROGRAMMING EXPERIENCE .....

LANGUAGES

MATLAB, C, C++, MAPLE, MATHEMATICA

OPERATING SYSTEMS

LINUX, WINDOWS XP, VMS

REFERENCES .....

Prof. Mingzhou Ding. Center for Complex Systems and Brain Sciences, 777 Glades Rd, Boca Raton, FL, 33431. Phone: (561)2972324; E-mail: ding@walt.ccs.fau.edu

Prof. Steven Bressler. Center for Complex Systems and Brain Sciences, 777 Glades Rd, Boca Raton, FL, 33431. Phone: (561)2972322; E-mail: bressler@walt.ccs.fau.edu

Prof. Dawei Dong. Center for Complex Systems and Brain Sciences, 777 Glades Rd, BocaRaton, FL, 33431. Phone: (561)2972326; E-mail: dawei@dove.ccs.fau.edu

Prof. J.A. Scott Kelso. Center for Complex Systems and Brain Sciences, 777 Glades Rd, Boca Raton, FL, 33431. Phone: (561)2972229; E-mail: kelso@walt.ccs.fau.edu