Fundamental Problems in Granger Causality Analysis of Neuroscience Data

by

Patrick A. Stokes

B.S. Biomedical Engineering and B.A. Economics,
University of Texas at Austin (2006)

Submitted to the Harvard-MIT Program in Health Sciences and Technology
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Medical Engineering and Medical Physics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2015

Signature redacted

Author .....
Harvard-MIT Program in Health Sciences and Technology
Signature redacted December 3, 2014

Certified by .....
Patrick L. Purdon, PhD
Assistant Professor of Anaesthesia, Harvard Medical School and
Massachusetts General Hospital
Signature redacted Thesis Supervisor

Accepted by .....
Emery N. Brown, MD, PhD
Director, Harvard-MIT Program in Health Sciences and Technology
Professor of Computational Neuroscience and Health Sciences and Technology

Fundamental Problems in Granger Causality Analysis of Neuroscience Data

by

Patrick A. Stokes

Submitted to the Harvard-MIT Program in Health Sciences and Technology
on December 3, 2014, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Medical Engineering and Medical Physics

Abstract

Granger causality methods analyze the flow of information between time series. The Geweke measure of Granger causality (GG-causality) has been widely applied in neuroscience because its frequency-domain and conditional forms appear well-suited to highly-multivariate oscillatory data. In this work, I analyze the statistical and structural properties of GG-causality in the context of neuroscience data analysis.

1. I analyze simulated examples and derive analytical expressions to demonstrate how computational problems arise in current methods of estimating conditional GG-causality. I show that the use of separate full and reduced models in the computation leads to either large biases or large uncertainties in the causality estimates, and high sensitivity to uncertainties in model parameter estimates, producing spurious peaks, valleys, and even negative values in the frequency domain.

2. I formulate a method of correctly computing GG-causality that resolves the above computational problems.

3. I analyze how generative system properties and frequency structure map into GG-causality to demonstrate deeper conceptual pitfalls:

   (a) I use simulated examples and derive analytical expressions to show that GG-causality is independent of the receiver dynamics, particularly the magnitude of response, which is counter-intuitive to physical notions of causality.

   (b) Overall, GG-causality combines transmitter and channel dynamics in a way that cannot be disentangled without evaluating the component dynamics of the full model estimate.

4. I discuss relevant concepts from causality analyses in other fields to better place causality analysis in a modeling and system identification framework.

The computational uncertainties in GG-causality estimates make the interpretation of frequency-domain structure highly problematic. Even if these computational issues are overcome, correct interpretation of the GG-causality values is still challenging and could be easily misinterpreted without careful consideration of the component dynamics of the full model estimate. Through this work, I provide conceptual clarification of GG-causality and place it in the broader framework of modeling and system analysis, which may enable investigators to better assess the utility and interpretation of such methods.

Thesis Supervisor: Patrick L. Purdon, PhD
Title: Assistant Professor of Anaesthesia, Harvard Medical School and Massachusetts General Hospital

Acknowledgments

I would like to thank my advisers, Patrick Purdon and Emery Brown, as well as my other committee members, Syd Cash and Uri Eden. I would also like to thank my previous mentors, Steve Kornguth and Mia Markey.

NIH Bioinformatics and Genomics Training Grant: T32 HG002295.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 The Appeal of Causality Analysis in Neuroscience

Neuroscience seeks to understand the mechanisms governing physiological states and processes within the brain. There are many sources of functional neural data covering various spatial and temporal scales, such as electroencephalography (EEG), magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), local field potentials (LFP). The objective is to quantitatively model, from the cellular to whole-brain level, functional neural systems during sensory stimuli, motor tasks, behavioral states, cognitive processes, and pathological activity that generate such data, to identify the component areas and to understand how the behavior of each component influences the behavior of the others.

Currently, most whole-brain level investigations merely identify regions of activity or, at best, analyze correlations between active regions. Causality analysis is set of system analysis methods which have recently become widely applied in neuroscience and that aims to go beyond correlations. Causality analysis purports to identify directed connections between components, assessing the directness of such influence, and/or characterizing the functional properties of such interactions. However, due to the wide variety of perspectives, there are many aspects of system behavior that one may consider causal. Thus one may ask: *What characterization of functional properties is intended by causality statements in neuroscience?*

One characterization of particular interest in neuroscience is the frequency-domain behavior. Oscillations are a ubiquitous feature of neurophysiological systems and neuroscience data. They are thought to constrain and organize neural activity within and between functional networks across a wide range of temporal and spatial scales. Oscillations at specific frequencies have been associated with different states of arousal [44], as well as different sensory [13], and cognitive processes [2, 10]. The prevalence of these neural oscillations, as well as their frequency-specific functional associations, suggest that frequency-domain properties are of interest in such statements of causality. Consequently, one motivator for the adoption of causality analyses in neuroscience has been the formulation of frequency-domain causality methods.

One approach to causality analysis is Granger causality, a statistical tool developed to analyze the flow of information between time series. Granger causality has been applied in neuroscience to many sources of data, including EEG, MEG, fMRI, and LFP. These studies have investigated functional neural systems across the many scales of organization, from the cellular level [36, 37, 25] up to whole-brain network activity [19], under a range of conditions, including sensory stimuli [13, 47, 52], varying levels of consciousness [4, 8, 50], cognitive tasks [28], and pathological states [14, 51]. In these analyses, the time series data are interpreted to reflect neural activity from particular sources, and Granger causality is used to characterize and quantify the directionality and directness of influence between sources.

A set of measures of Granger causality was proposed by Geweke [26, 27]. These include time-domain and frequency-domain measures. While the causality measures of [26] are unconditional (bivariate), those of [27] are conditional (multivariate), allowing inclusion of additional time series and offering a means to distinguish direct influences from indirect influences between system components within a larger network. Such multivariate methods are further appealing in neuroscience where data are often highly multivariate. Hence, the Granger-Geweke approach seems to offer neuroscientists precisely what they want—an assessment of direct frequency-dependent functional influence between time series—in a straightforward extension of standard time series techniques. Current, widely-applied tools for analyzing neuroscience data

are based on this Granger-Geweke (GG) causality method [12, 49, 7]. Subsequently, many alternative frequency-domain causality measures have been proposed, such as the directed transfer function (DTF) [34, 35], the partial coherence, the direct directed transfer function (dDTF) [39, 38], and the partial directed coherence (PDC) [5].

1.2 Example Application: Epilepsy

A common application of causality analysis is epilepsy. One goal of such applications is to improve localization of seizure-onset regions, to guide surgical intervention while minimizing resection and thereby improve efficacy and reduce morbidity. Epstein et al. [20] applied GG-causality to data from 10 retrospective patients and 2 prospective patients with seizures refractory to medication. Intracranial EEG (iEEG) data was obtained for each patient, from various combinations of depth electrodes, subdural grids, and/or strip electrodes. Sliding windows of frequency-domain GG-causality were estimated for data before and after seizure onset and compared with the results of standard visual analysis to determine if the findings were consistent with the surgical intervention and post-operative status.

Figure 1-1, reproduced from [20], shows causality estimates for two retrospective patients. The top plot shows the iEEG time series for retrospective patient 1. Visual analysis showed first ictal activity at 90 Hz in a strip electrode over the parahippocampal gyrus (spH5) and 65 Hz in depth electrode in the anterior hippocampal depth electrode (dAH1), suggesting spH5 as the seizure onset zone. The middle two plots show the frequency-domain GG-causality between spH5 and dAH1 for the first and second halves of the time series. The causality estimates suggest that dAH1 is the stronger source, even prior to ictal activity. The bottom plot shows the iEEG time series for retrospective patient 3 with GG-causality estimates overlaid and an inset showing location of electrode placement—a 10-contact superior parietal lobe grid and a 64-contact frontal-parietosuperior temporal grid. Again, the earliest significant causality (black arrow) occurs prior to the onset suggested by visual analysis of the

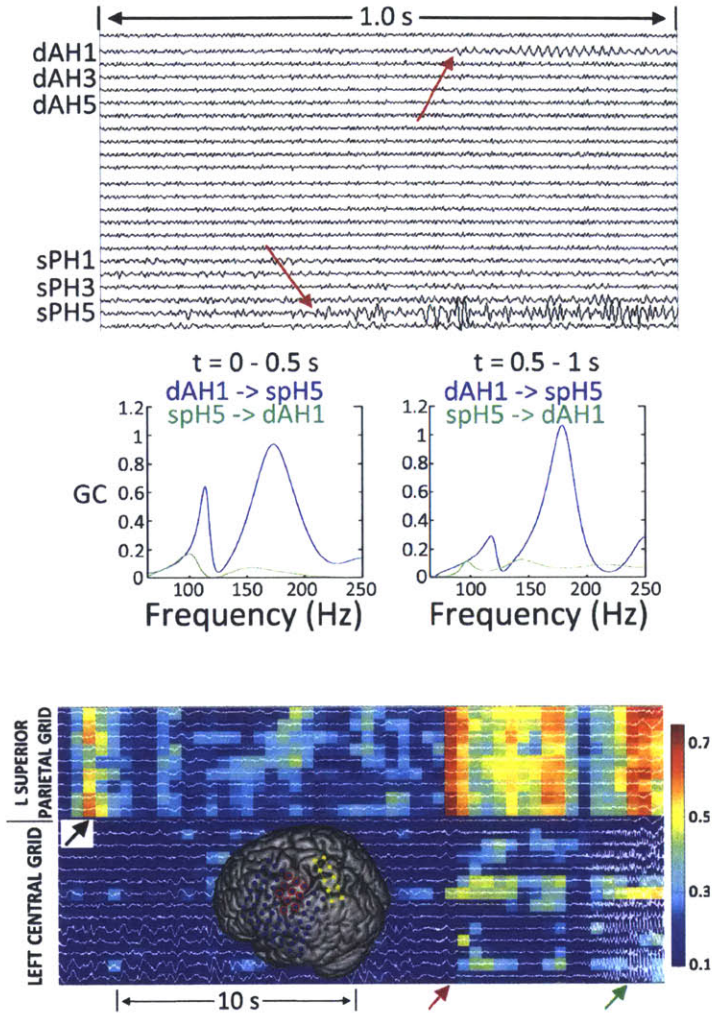

Figure 1-1: Causality estimates for two epilepsy patients. Reprinted from [20] by permission of publisher John Wiley & Sons, Inc.

iEEG (red arrow).

For the prospective patients studied in this paper, the GG-causality was used to exclude locations from resection that were otherwise indicated by prominent, early appearance in the visual analysis. Epstein et al. suggest that the retrospective and prospective surgical interventions and outcomes are generally consistent with their GG-causality estimates, though they acknowledge the limits of such conclusions, given the small number of randomly selected cases. They also acknowledge that it is not possible to determine if the positive outcomes represent removal of the seizure-onset zone or simply disruption of the propagation network. The interpretation of GG-causality in this analysis is that larger causality values indicate stronger influence or, equivalently, greater relative importance in determining the resulting system behavior, i.e. the high frequency oscillations (HFO) of seizure. The medical interest in and implications for this application of causality analysis is to guide surgical treatment. Therefore, it is crucial to correctly identify sources for inclusion and exclusion during resection.

The interpretation of GG-causality indicated in this example, that time-domain causality estimes quantify total strength of influence and that frequency-domain causality estimates quantify influence at specific frequencies, is quite prevalent, whether explicitly stated or not. This leads naturally to the question: *How does GG-causality relate to the functional properties of the generative system?*

1.3 Example Application: Time-Varying GG-Causality

Another avenue of causality investigations has been the extension of causality methods to data from time-varying processes. As is the case in the epilepsy application above, the time series involved are non-stationary. Also, due to the nature of seizure activity, it is also not possible to obtain an ensemble of replicated trials. An original aim of this thesis, and the impetus for the work presented, was to expand the GG-causality measures to allow for time-varying causality analysis in the frequency domain of single experimental realizations, and to apply the method to data for which the generative

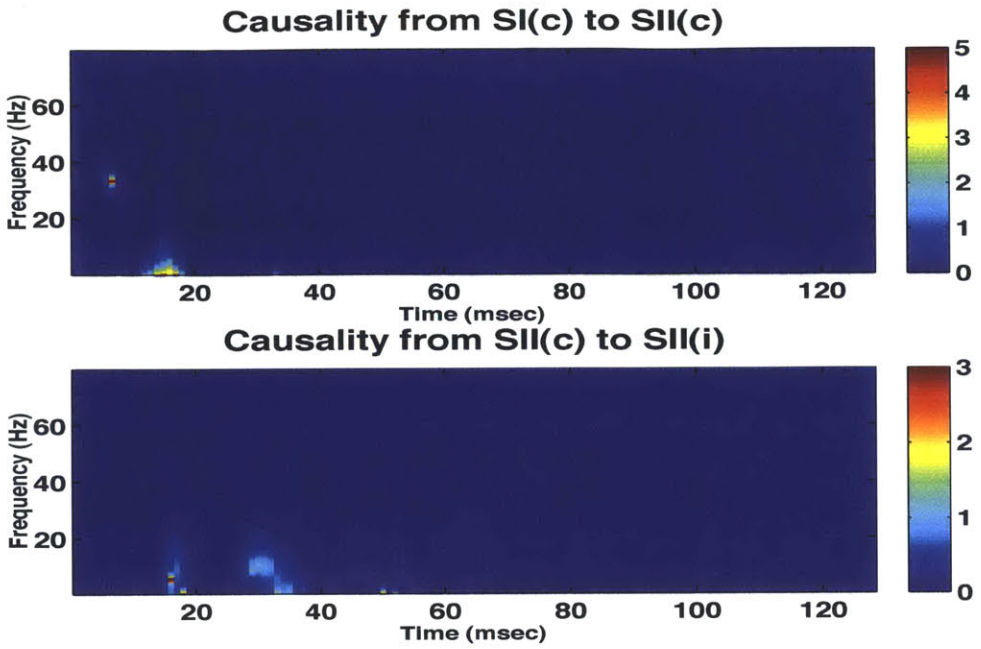

Figure 1-2: Time-varying, frequency-domain GG-causality estimates for a median nerve stimulation experiment. The top plot shows the causality from primary somatosensory cortex (SIc) to secondary somatosensory cortex (SIIc) on the side contralateral to the stimulus. The bottom plot shows the causality from contralateral secondary somatosensory cortex to ipsilateral secondary somatosensory cortex (SIIi). (Data courtesy of Tommi Raij and Matti Hamalainen, MGH Martinos Center for Biomedical Imaging.)

process has time-varying properties and/or repeated samples of the process are not obtainable. Some preliminary results from those early investigations are shown in the following figures.

Figure 1-2 shows time-varying, conditional frequency-domain causality estimates for a median nerve stimulation experiment. Two physiologically plausible points of positive causality are seen–a 35 Hz interaction from contralateral primary somatosensory to contralateral secondary somatosensory followed by a 10 Hz interaction between secondary somatosensory regions, from contralateral to ipsilateral. However, the identified causality is not particularly prominent, and most of the time-frequency plots are actually set to zero because the causality was estimated to be negative.

To assess the difficulty with the time-varying causality estimation, a simulated

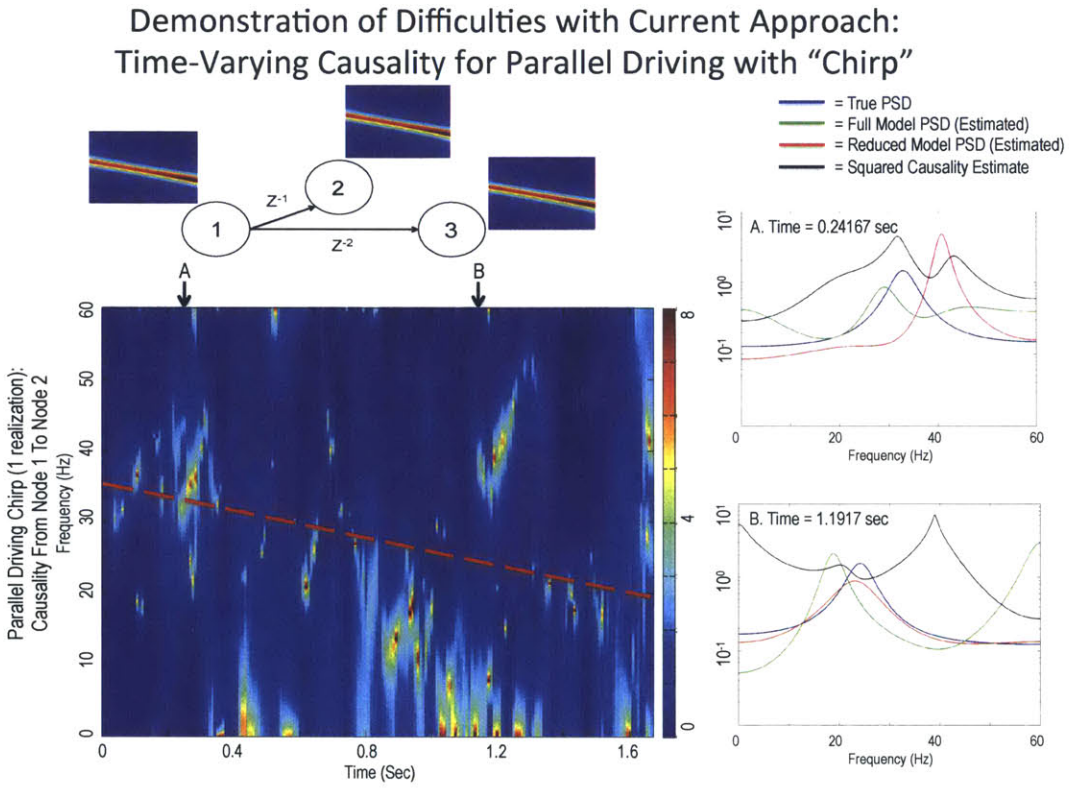

Figure 1-3: Time-varying causality estimates for a simulated example.

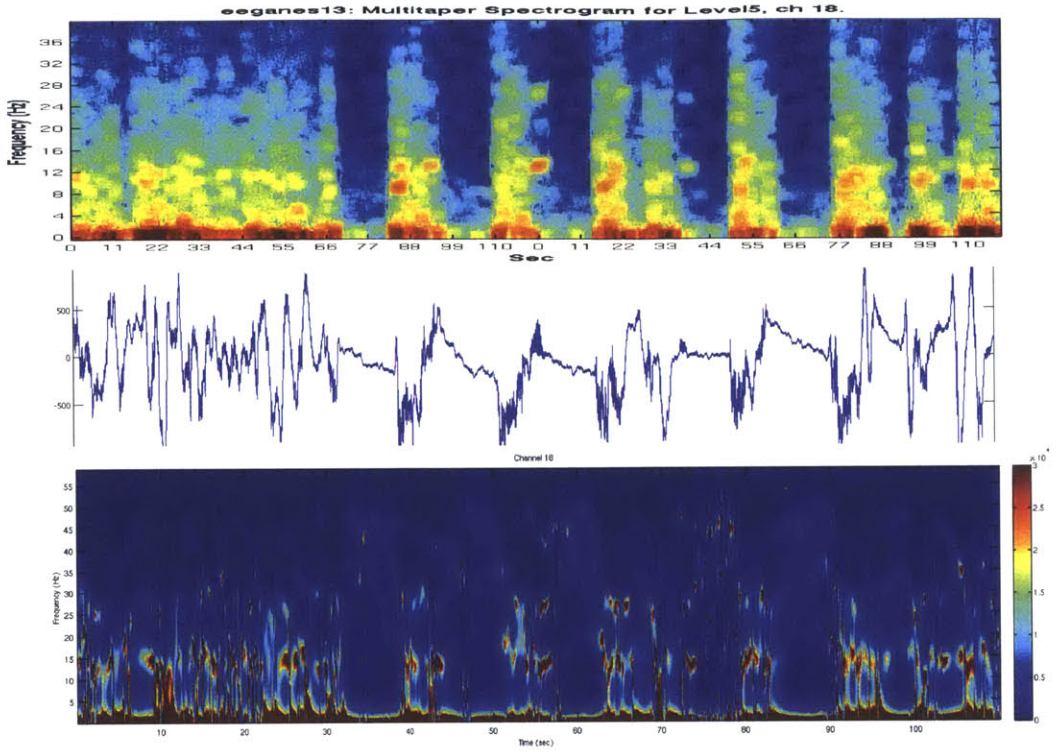

Figure 1-4: Time-varying parametric spectral estimates for EEG electrodes burst suppression.

example was analyzed. Figure 1-3 shows time-varying, frequency-domain causality estimates for a cause-effect pair of a parallel-delay driving example with a spectral chirp. The estimates are noticeably noisy, despite reasonable tracking of the spectra by the both the full and reduced models.

For a further diagnostic, the time-varying parametric spectral estimation was examined. Figure 1-4 shows time-varying parametric spectral estimates for EEG recordings of burst suppression during anesthesia. The estimates generally agree with those of the multi-taper estimation, with low-frequency power and a 10-15 Hz band, which appear during bursting. However, the parametric estimates are noisier from time-point to time-point, showing high frequency ripples, and a slight shift in the 10 Hz peak to 15 Hz.

These results revealed some difficulties that arise generally in GG-causality analysis: 1) the occurrence of negative values, 2) the noisiness of the estimates, notably

"ripples" from frequency-domain peaks and valleys, and 3) the dependence on the selected model order, as the location and number of peaks can vary wildly.

Further comparison with other causality approaches showed similar difficulties but also large disagreement between the estimated values of different causality measures. As there is no established "ground truth causality" against which to evaluate the results, the question arose *what is the true causality when the generative system is within the vector autoregressive model class underlying the GG-causality approach?*

1.4 The Problem

Many limitations of GG-causality are well-known, including the requirements that the generative system be approximately linear, stationary, and time-invariant. Application of the GG-approach to more general processes, such as neural spiking data [37, 25], continuous-time processes [43], and systems with exogenous inputs and latent variables [19, 30], has been shown to produce results inconsistent with the known functional structure of the underlying system. These examples illustrate the perils of applying GG-causality in situations where the generative system is poorly approximated by the VAR model class. Other problems, such as negative causality estimates [16], have been observed even when the generative system belongs to the finite order VAR model class. Together, these problems raise several important questions for neuroscientists interested in using GG-methods: Under what circumstances are these methods appropriate? How reliably do these methods recover the functional structure underlying the observed data? What do frequency-domain causality measures tell us about the functional properties of oscillatory neural systems?

Thus, in the course of applying GG-causality, several problems were encountered, both computational and conceptual, which can produce spurious results and be highly misleading, particularly in the neuroscience context. Consequently, this work became a demonstration and analysis of these problems in the hopes of providing clarification for those interested in using Granger causality and other causality analysis approaches.

1.5 Specific Contributions

In this work, I analyze GG-causality methods to help address the questions raised above. The specific contributions of this work are as follows:

1. I use a combination of simulation and mathematical analysis to demonstrate and analyze computational problems that arise in estimation of conditional GG-causality. I show how the use of separate full and reduced model fits leads to a peculiar bias-variance trade-off in model order, such that use of the true model order results in biased causality estimates, while increasing the order increases the variance of the estimates and produces spurious peaks and valleys in the frequency-domain. I also show how the separate model fits also result in high sensitivity of the GG-causality estimates to uncertainties in the estimated model parameters, leading to additional spurious peaks and valleys and even negative values. The uncertainties introduced by these computational problems have thus far masked more fundamental interpretational problems that also undermine GG-causality analysis.

2. I present the solution to the computational problems analyzed in Contribution 1. I show that by putting the full VAR model in state-space form, standard spectral factorization results can be used to obtain the desired reduced model, from which unbiased GG-causality estimates of minimal variance can be computed.

3. I use simulation and mathematical analysis to analyze how the functional properties of the underlying system map into the GG-causality measures. I show that the GG-causality measures are generally independent of the dynamics of the receiver node (effect component), but remain a function of the transmitter and channel dynamics. I clarify how these properties of GG-causality can lead to misinterpretation, especially in neuroscience, as in the epilepsy application in [20]. These results lead to a more general discussion of how GG-causality relates to other proposed causality measures, and the potential role of causality analysis in neuroscience applications.

In this work, the analysis is limited to simulations of finite order VAR systems to ensure that the problems demonstrated are not due to inaccurate modeling of the generative process, but rather are inherent to GG-causality. Additionally, I focus my analysis on the Geweke measures—unconditional and conditional, time-domain and frequency-domain—of Granger causality for several reasons. The Geweke measures are some of the more commonly used causality measures, particularly the unconditional and time-domain cases, and are widely available in MATLAB toolboxes [49, 7]. Furthermore, the computational problems analyzed in Contribution 1 and resolved in Contribution 2 are partly what have limited the use of the conditional case and have motivated the proposal of related alternative causality measures. Lastly, the Geweke frequency-domain measures largely motivated the adoption of Granger causality analysis and related methods in the neurosciences.

A broader aim of this work is to clarify, on several levels, the current status of causality analysis in neuroscience. There is an unacknowledged ambiguity in the pursuit of neuroscientific causality. This has left ample opportunity for misunderstanding by both investigators and readers of previously published causality analyses. The analyses and results of my work and the subsequent discussion provide important conceptual clarification of GG-causality, and causality analyses in general, by placing it in the broader context of modeling and system identification. In this framework, there is a clear relationship between the various VAR-based causality measures. There are also parallels between causality arguments in neuroscience and how causality questions have been addressed in other fields, like econometric time series, classical regression and statistical inference, and probabilistic reasoning. This comparison is informative of how different modeling perpectives and objectives alter the way such questions are answered. In expounding on these topics, I hope to enable readers to better assess the utility of causality analysis methods, to gauge their own interest in applying such methods, and to properly interpret results.

1.6 Summary

This work is organized as follows:

- In Chapter 2, I review Granger causality and the Geweke measures that are the focus of this work. I also briefly review several other causality measures used in neuroscience investigations and highlight some of the many applications.

- In Chapter 3, I demonstrate and analyze the problems that arise when computing the conditional Geweke causality measures using separate model fits. I show how these problems introduce biases and uncertainties that can mask deeper issues with interpretation and modeling.

- In Chapter 4, I present the solution to the problem of computing conditional GG-causality by using the correct spectral factorization of the state-space form of the full model. The improved results are demonstrated for a simulated example.

- In Chapter 5, I analyze how the properties of the generative system map into the GG-causality measures. I show using a simulated example that interpretational problems undermine GG-causality analyses and can be highly misleading, especially in neuroscience applications.

- In Chapter 6, I discuss more broadly the role of causality analysis in neuroscience by placing it in a modeling and system identification framework and relating it to causality analysis in other fields.

Chapter 2

Background

2.1 Granger Causality

Granger causality developed in the field of econometric time series analysis. Granger [29] formulated a statistical definition of causality from the premises that (i) a cause occurs before its effect and that (ii) knowledge of a cause improves prediction of its effect. Under this framework, a time series $\{x_{i,t}\}$ is Granger-causal of another time series $\{x_{j,t}\}$ if inclusion of the history of x_i improves prediction of x_j over knowledge of the history of x_j alone. Specifically, this is quantified by comparing the prediction error variances of the one-step linear predictor, $\hat{x}_{j,t}$, under two different conditions: one where the full histories of all time series are used for prediction, and another where the putatively causal time series is omitted from the set of predictive time series. Thus, $\{x_{i,t}\}$ Granger-causes $\{x_{j,t}\}$ if

$$\mathrm{var}\left(x_{j,t} - \hat{x}_{j,t} \mid x_{j,0:t-1}\right) > \mathrm{var}\left(x_{j,t} - \hat{x}_{j,t} \mid x_{j,0:t-1}, x_{i,0:t-1}\right).$$

Granger causality can be generalized to multi-step predictions, higher-order moments of the prediction distributions, and alternative predictors.

In practice, the above linear predictors are limited to finite order vector autoregressive (VAR) models, and Granger causality is assessed by comparing the prediction error variances from separate VAR models, with one model including all components

of the vector time series data, termed the full model, and a second including only a subset of the components, termed the reduced model. To investigate the causality from $\{x_{i,t}\}$ to $\{x_{j,t}\}$, let

$$\begin{bmatrix} x_{j,t} \\ \\ x_{i,t} \end{bmatrix} = \sum_{p=1}^{P} \begin{bmatrix} A_{j,j}^{f}(p) & A_{j,i}^{f}(p) \\ \\ A_{i,j}^{f}(p) & A_{i,i}^{f}(p) \end{bmatrix} \begin{bmatrix} x_{j,t-p} \\ \\ x_{i,t-p} \end{bmatrix} + \begin{bmatrix} w_{j,t}^{f} \\ \\ w_{i,t}^{f} \end{bmatrix} \tag{2.1}$$

be the full VAR(P) model of all time series components, where the superscript f is used to denote the full model. This model may be written more compactly as $x_t^f = \sum_{p=1}^{P} A^f(p)\, x_{t-p} + w_t^f$. The noise processes $\{w_{j,t}^f\}$ and $\{w_{i,t}^f\}$ are zero-mean and temporally uncorrelated with covariance $E\left[w_{t_1}^f \left(w_{t_2}^f\right)^T\right] = \Sigma^f \delta_{t_1 - t_2}$. Thus, the full one-step predictor of $x_{j,t}$ in the above causality definition is $\hat{x}_{j,t}^f = \sum_{p=1}^{P} \sum_{m \in \{i,j\}} A_{j,m}^f(p)\, x_{m,t-p}$. Similarly, let

$$x_{j,t} = \sum_{p=1}^{P} A^{r(i)}(p)\, x_{j,t-p} + w_{j,t}^{r(i)}$$

be the reduced VAR(P) model of the x_j (putative effect) components of the time series, omitting the x_i (putative cause) components. The superscript $r(i)$ is used to denote this reduced model formed by omitting x_i. The noise process $\{w_{j,t}^{r(i)}\}$ is zero-mean and temporally uncorrelated with covariance $\Sigma^{r(i)}$. The reduced one-step predictor of $x_{j,t}$ is thus $\hat{x}_{j,t}^{r(i)} = \sum_{p=1}^{P} A^{r(i)}(p)\, x_{j,t-p}$.

2.2 Geweke Time-Domain Causality Measures

Building on Granger's definition, Geweke [26] defined a measure of directed causality (what he referred to as linear feedback) from $\{x_{i,t}\}$ to $\{x_{j,t}\}$ to be

$$\mathcal{F}_{x_i \longrightarrow x_j} = \ln \frac{\left|\mathrm{var}\left(x_{j,t} - \hat{x}_{j,t}^{r(i)}\right)\right|}{\left|\mathrm{var}\left(x_{j,t} - \hat{x}_{j,t}^{f}\right)\right|} = \ln \frac{\left|\Sigma_{j,j}^{r(i)}\right|}{\left|\Sigma_{j,j}^{f}\right|},$$

where $|\Sigma| = \sqrt{\det \Sigma' \Sigma}$. Geweke [27] expanded the previous definition of unconditional (bivariate) causality to include conditional time series components $\{x_{k,t}\}$, making it possible to distinguish between direct and indirect influences. For example, consider three time series $\{x_{i,t}\}$, $\{x_{j,t}\}$, and $\{x_{k,t}\}$. By conditioning on x_k, it is possible to distinguish between direct influences between x_i and x_j, as opposed to indirect influences that are mediated by x_k. The conditional measure, $\mathcal{F}_{x_i \longrightarrow x_j | x_k}$, has a form analogous to the unconditional case, except that the predictors, $\hat{x}_{j,t}^{f}$ and $\hat{x}_{j,t}^{r(i)}$, and associated prediction-error variances, $\Sigma_{j,j}^{f}$ and $\Sigma_{j,j}^{r(i)}$, all incorporate the conditional time series $\{x_{k,t}\}$.

These time-domain causality measures have a number of important properties. First, they are theoretically non-negative, equaling zero in the case of non-causality. Second, the total linear dependence between two time series can be represented as the sum of the directed causalities and an instantaneous causality (See the Appendix 2.4.1 for details). Finally, these time-domain causality measures can further be decomposed by frequency, providing a frequency-domain measure of causality.

2.3 Geweke Frequency-Domain Causality Measures

The above time-domain measure of causality affords a spectral decomposition, which allowed Geweke [26] to also define an unconditional frequency-domain measure of causality. Let $X(\lambda) = H^f(\lambda) W^f(\lambda)$ be the frequency-domain representation of the moving-average (MA) form of the full model in Eqn. 2.1. $X(\lambda)$ and $W^f(\lambda)$ are the Fourier representations of the vector time series $\{x_t\}$ and noise process $\{w_t^{(f)}\}$, respectively, and $H^f(\lambda)$ is the transfer function given by

$$H^f(\lambda) = \left(I - \sum_{p=1}^{P} A^f(p) e^{-ip\lambda} \right)^{-1}.$$

As alluded to above, the model can contain instantaneously causal components. The frequency-domain definition requires removal of the instantaneous causality components by transforming the system with a rotation matrix, as described in Ap-

pendix 2.4.2 and [26]. For clarity, this rotation is omitted from the present overview of frequency-domain causality, but the transformation was fully implemented in the computational studies that follow. The spectrum of $\{x_{j,t}\}$ is then

$$S_{x_j,x_j}(\lambda) = H^f_{j,j}(\lambda)\,\Sigma^f_{j,j}H^{f*}_{j,j}(\lambda) + H^f_{j,i}(\lambda)\,\Sigma^f_{i,i}H^{f*}_{j,i}(\lambda),$$

where $*$ denotes conjugate transpose. The first term is the component of the spectrum of x_j due to its own input noise process, while the second term represents the components introduced by the x_i time series. The unconditional frequency-domain causality from $\{x_{i,t}\}$ to $\{x_{j,t}\}$ at frequency λ is defined as

$$\begin{aligned} f_{x_i\longrightarrow x_j}(\lambda) &= \ln \frac{|S_{x_j,x_j}(\lambda)|}{|H^f_{j,j}(\lambda)\,\Sigma^f_{j,j}H^{f*}_{j,j}(\lambda)|} \\ &= \ln \frac{|H^f_{j,j}(\lambda)\,\Sigma^f_{j,j}H^{f*}_{j,j}(\lambda) + H^f_{j,i}(\lambda)\,\Sigma^f_{i,i}H^{f*}_{j,i}(\lambda)|}{|H^f_{j,j}(\lambda)\,\Sigma^f_{j,j}H^{f*}_{j,j}(\lambda)|}. \end{aligned} \tag{2.2}$$

If $\{x_{i,t}\}$ does not contribute to the spectrum of $\{x_{j,t}\}$ at a given frequency, the second term in the numerator of Eqn. 2.2 is zero at that frequency, resulting in zero causality. Thus, the unconditional frequency-domain causality reflects the components of the spectrum originating from the input noise of the putatively causal time series.

As with the time-domain measure, Geweke [27] expanded the frequency-domain measure to include conditional time series. And like the conditional time-domain measure, the conditional frequency-domain definition requires separate full and reduced models. Let $H^f(\lambda)$ and Σ^f be the system function and noise covariance of the full model, and let $H^{r(i)}(\lambda)$ and $\Sigma^{r(i)}$ be the system function and noise covariance of the reduced model. As in the unconditional case, instantaneous causality components must be removed, this time from each model, as described in Appendix 2.4.2 and [27]. Again, for clarity, this rotation is omitted from the equations below, but fully implemented in the computational studies that follow.

The derivation for the conditional form of frequency-domain causality relies on the time-domain equivalence demonstrated by Geweke [27], $\mathcal{F}_{x_i\longrightarrow x_j|x_k} = \mathcal{F}_{x_i w_k^{r(i)}\longrightarrow w_j^{r(i)}}$.

The conditional frequency-domain causality is then defined by

$$f_{x_i \longrightarrow x_j | x_k}(\lambda) = f_{x_i w_k^{r(i)} \longrightarrow w_j^{r(i)}}(\lambda).$$

Thus, original VAR model must be rewritten in terms of the time series $\{w_{j,t}^{r(i)}\}$, $\{x_{i,t}\}$, and $\{w_{k,t}^{r(i)}\}$. Geweke [27] did this by cascading the full model with the inverse of an augmented form of the reduced model,

$$\begin{bmatrix} W_j^{r(i)}(\lambda) \\ X_i(\lambda) \\ W_k^{r(i)}(\lambda) \end{bmatrix} = \begin{bmatrix} H_{j,j}^{r(i)}(\lambda) & 0 & H_{j,k}^{r(i)}(\lambda) \\ 0 & I & 0 \\ H_{k,j}^{r(i)}(\lambda) & 0 & H_{k,k}^{r(i)}(\lambda) \end{bmatrix}^{-1} \begin{bmatrix} H_{j,j}^{f}(\lambda) & H_{j,i}^{f}(\lambda) & H_{j,k}^{f}(\lambda) \\ H_{i,j}^{f}(\lambda) & H_{i,i}^{f}(\lambda) & H_{i,k}^{f}(\lambda) \\ H_{k,j}^{f}(\lambda) & H_{k,i}^{f}(\lambda) & H_{k,k}^{f}(\lambda) \end{bmatrix} \begin{bmatrix} W_j^{f}(\lambda) \\ W_i^{f}(\lambda) \\ W_k^{f}(\lambda) \end{bmatrix} \tag{2.3}$$

$$= \begin{bmatrix} G_{j,j}(\lambda) & G_{j,i}(\lambda) & G_{j,k}(\lambda) \\ G_{i,j}(\lambda) & G_{i,i}(\lambda) & G_{i,k}(\lambda) \\ G_{k,j}(\lambda) & G_{k,i}(\lambda) & G_{k,k}(\lambda) \end{bmatrix} \begin{bmatrix} W_j^{f}(\lambda) \\ W_i^{f}(\lambda) \\ W_k^{f}(\lambda) \end{bmatrix}. \tag{2.4}$$

This combined system relates the reduced model noise processes and the putatively causal time series to the full model noise process. It can be viewed as the frequency-domain representation of a VAR model of the time series $\{w_{j,t}^{r(i)}\}$, $\{x_{i,t}\}$, and $\{w_{k,t}^{r(i)}\}$. The spectra of $\{w_{j,t}^{r(i)}\}$ is then

$$S_{w_j^{r(i)}, w_j^{r(i)}}(\lambda) = G_{j,j}(\lambda)\,\Sigma_{j,j}^{f} G_{j,j}^{*}(\lambda) + G_{j,i}(\lambda)\,\Sigma_{i,i}^{f} G_{j,i}^{*}(\lambda) + G_{j,k}(\lambda)\,\Sigma_{k,k}^{f} G_{j,k}^{*}(\lambda). \tag{2.5}$$

Hence, the conditional frequency-domain causality is

$$f_{x_i \longrightarrow x_j | x_k}(\lambda) = \ln \frac{|S_{w_j^{r(i)}, w_j^{r(i)}}(\lambda)|}{|G_{j,j}(\lambda)\,\Sigma_{j,j}^{f} G_{j,j}^{*}(\lambda)|} = \ln \frac{|\Sigma_{j,j}^{r(i)}|}{|G_{j,j}(\lambda)\,\Sigma_{j,j}^{f} G_{j,j}^{*}(\lambda)|}.$$

The last equality follows from the assumed whiteness of $\{w_{j,t}^{r(i)}\}$ with covariance $\Sigma_{j,j}^{r(i)}$.

This equation is very similar in form to the expression for unconditional frequency-domain causality Eqn. 2.2, suggesting an analogous interpretation. However, the

transformations detailed in Eqn. 2.3 belie a more nuanced interpretation. While the unconditional frequency-domain causality depends on the component of the effect spectrum not due to its own input noise, the conditional frequency-domain causality is the component of the reduced model uncertainty not due to the corresponding full model input noise.

2.4 Appendix: Causality and VAR Miscellany

2.4.1 Instantaneous Causality and Total Linear Dependence

In addition to the unconditional and conditional directed causality measures, $\mathcal{F}_{i\longrightarrow j}$ and $\mathcal{F}_{i\longrightarrow j|k}$, Geweke [26] and [27] also defined associated (undirected) instantaneous causality measures. The unconditional instantaneous causality is

$$\mathcal{F}_{i\cdot j} = \ln \frac{|\text{var}\left(x_{j,t} - \hat{x}^f_{j,t}\right)| \cdot |\text{var}\left(x_{i,t} - \hat{x}^f_{i,t}\right)|}{|\text{var}\left(x_{\cdot,t} - \hat{x}^f_{\cdot,t}\right)|} = \ln \frac{|\Sigma^f_{j,j}||\Sigma^f_{i,i}|}{|\Sigma^f_{\cdot,\cdot}|},$$

where the subscript "$\cdot$" denotes all components of the system, i.e. $\{j, i\}$ in the unconditional case. The expression for the instantaneous causality in the conditional case, $\mathcal{F}_{i\cdot j|k}$, is identical, with the conditional components k included in the models.

An alternative measure quantifying the dependence between two time series, e.g. between $\{x_{i,t}\}$ and $\{x_{j,t}\}$, is the total linear dependence,

$$\mathcal{F}_{i,j} = \ln \frac{|\text{var}\left(x_{j,t} - \hat{x}^{r(i)}_{j,t}\right)| \cdot |\text{var}\left(x_{i,t} - \hat{x}^{r(j)}_{i,t}\right)|}{|\text{var}\left(x_{\cdot,t} - \hat{x}^f_{\cdot,t}\right)|} = \ln \frac{|\Sigma^{r(i)}_{j,j}||\Sigma^{r(i)}_{i,i}|}{|\Sigma^f_{\cdot,\cdot}|}.$$

Hence, one additional motivation of the definition of directed and instantaneous causality of Geweke [26] is that they comprise a time-domain decomposition of the total linear dependence, such that

$$\mathcal{F}_{i,j} = \mathcal{F}_{i\longrightarrow j} + \mathcal{F}_{j\longrightarrow i} + \mathcal{F}_{i\cdot j}.$$

The expression for the total linear dependence in the conditional case, $\mathcal{F}_{i,j|k}$, is identical, such that the conditional definitions of directed and instantaneous causality of [27] similarly comprise a time-domain decomposition,

$$\mathcal{F}_{i,j|k} = \mathcal{F}_{i \longrightarrow j|k} + \mathcal{F}_{j \longrightarrow i|k} + \mathcal{F}_{i \cdot j|k}.$$

2.4.2 Removal of Instantaneous Causality for Frequency-Domain Computation

The definition of unconditional frequency-domain causality requires removal of the instantaneous component by transforming the system function and input noise process. Given the frequency-domain representation of the full model

$$X(\lambda) = H^f(\lambda) W^f(\lambda)$$

with output spectrum

$$S(\lambda) = H^f(\lambda) \Sigma^f \left(H^f(\lambda)\right)^*,$$

post-multiplication of the system function by D^{-1} and pre-multiplication of the noise process by D, where

$$D = \begin{bmatrix} \mathrm{I} & 0 \\ -\Sigma^f_{ij} \left(\Sigma^f_{jj}\right)^{-1} & \mathrm{I} \end{bmatrix},$$

leaves the output spectrum invariant, but rotates the system to a model with diagonal input noise covariance $\tilde{\Sigma}^f = D\Sigma^f D^T$ and transformed system function $\tilde{H}^f(\lambda) = H^f(\lambda) D^{-1}$. Here, as in the rest of the text, the following ordering convention is used: if one is investigating the causality from x_i to x_j, the effect components are ordered before the causal components, i.e. $x_t = \begin{bmatrix} x^T_{j,t} & x^T_{i,t} \end{bmatrix}^T$.

In the conditional case, the frequency-domain causality computation requires removal of the instantaneous causality from both the full model and the reduced model. For the reduced model, the system function $H^{r(i)}(\lambda)$ and noise covariance $\Sigma^{r(i)}$ are

transformed, as above, to $\tilde{H}^{r(i)}(\lambda)$ and $\tilde{\Sigma}^{r(i)}$, using

$$D^{r(i)} = \begin{bmatrix} \mathrm{I} & 0 \\ -\Sigma_{kj}^{r(i)} \left(\Sigma_{jj}^{r(i)}\right)^{-1} & \mathrm{I} \end{bmatrix}.$$

Similarly, the system function $H^f(\lambda)$ and noise covariance matrix Σ^f of the full model are transformed to $\tilde{H}^f(\lambda)$ and $\tilde{\Sigma}^f$, but in this case, the transformation matrix for the full model (See the appendix of [16]) is given by

$$D^f = \begin{bmatrix} \mathrm{I} & 0 & 0 \\ 0 & \mathrm{I} & 0 \\ 0 & -\left(\Sigma_{ki}^f - \Sigma_{kj}^f \left(\Sigma_{jj}^f\right)^{-1} \Sigma_{jk}^f\right) \left(\Sigma_{ii}^f - \Sigma_{ij}^f \left(\Sigma_{jj}^f\right)^{-1} \Sigma_{ji}^f\right)^{-1} & \mathrm{I} \end{bmatrix}$$
$$\times \begin{bmatrix} \mathrm{I} & 0 & 0 \\ -\Sigma_{ij}^f \left(\Sigma_{jj}^f\right)^{-1} & \mathrm{I} & 0 \\ -\Sigma_{kj}^f \left(\Sigma_{jj}^f\right)^{-1} & 0 & \mathrm{I} \end{bmatrix}.$$

Note that the nature of the transformations now make the causality computation dependent on the order of the components in the vector time series. Again, the ordering convention for the rest of the text is followed. When investigating the causality from x_i to x_j conditional on x_k, the effect components are ordered before the causal components in both the full and reduced models and the conditional components are appended to the end of the full model, i.e. $\{j, i\}$ and $\{j, i, k\}$.

2.4.3 MA Matrix Recursions for Invertible VAR Systems

In the computation of GG-causality, the full and reduced models are assumed to be stable and invertible so that the transfer functions can be computed. The VAR(P) system

$$x_t - \sum_{p=1}^{P} A(p)\, x_{t-p} = w_t$$

is invertible if

$$\det\left(I - \sum_{p=1}^{P} A(p) z^p\right) \neq 0 \text{ for } |z| \leq 0.$$

The transfer function is then

$$H(\lambda) = \left(I - \sum_{p=1}^{P} A(p) e^{-ip\lambda}\right)^{-1}.$$

The transfer function is equivalently the Fourier transform of the MA matrices. In MA form, the system is

$$x_t = \sum_{p=0}^{\infty} \Phi(p) w_{t-p}.$$

Though they are not utilized in this work, for convenience the MA matrix recursions are included here. The MA matrices can be computed recursively from the AR matrices as follows [42]:

$$\Phi(0) = I_M$$

$$\Phi(p) = \sum_{j=1}^{p} \Phi(p-j) A(j) \text{ for } p = 1, 2, \ldots.$$

Chapter 3

Computational Problems

3.1 Consequences of the VAR Model Class on Conditional Causality Estimates

This chapter illustrates how the properties of the VAR model impact estimation of conditional GG-causality. It shows that the structure of VAR model class, in combination with the choice to use separate full and reduced models to estimate causality, introduces a peculiar bias-variance trade-off to the estimation of conditional frequency-domain GG-causality. In particular, the reduced model is used to fit a subset of a VAR process, which is, generally, outside the VAR model class. Consequently, there is a trade-off between using the true full model order, which forces the reduced-model representations to be biased, leading to biased causality, versus a higher model order, which better represents the reduced process, but overfits the full model, leading to high GG-causality variance, as well as spurious frequency-domain peaks and valleys. The separate model fits also introduce an additional sensitivity to parameter estimates, exacerbating the problem of spurious peaks and valleys, and even producing negative values.

Example 1: VAR(3) 3-Node Series System

This example analyzes time series generated by the VAR(3) system shown in Figure 3-1. In this system, node 1, $\{x_{1,t}\}$, resonates at frequency $f_1 = 40$ Hz and is

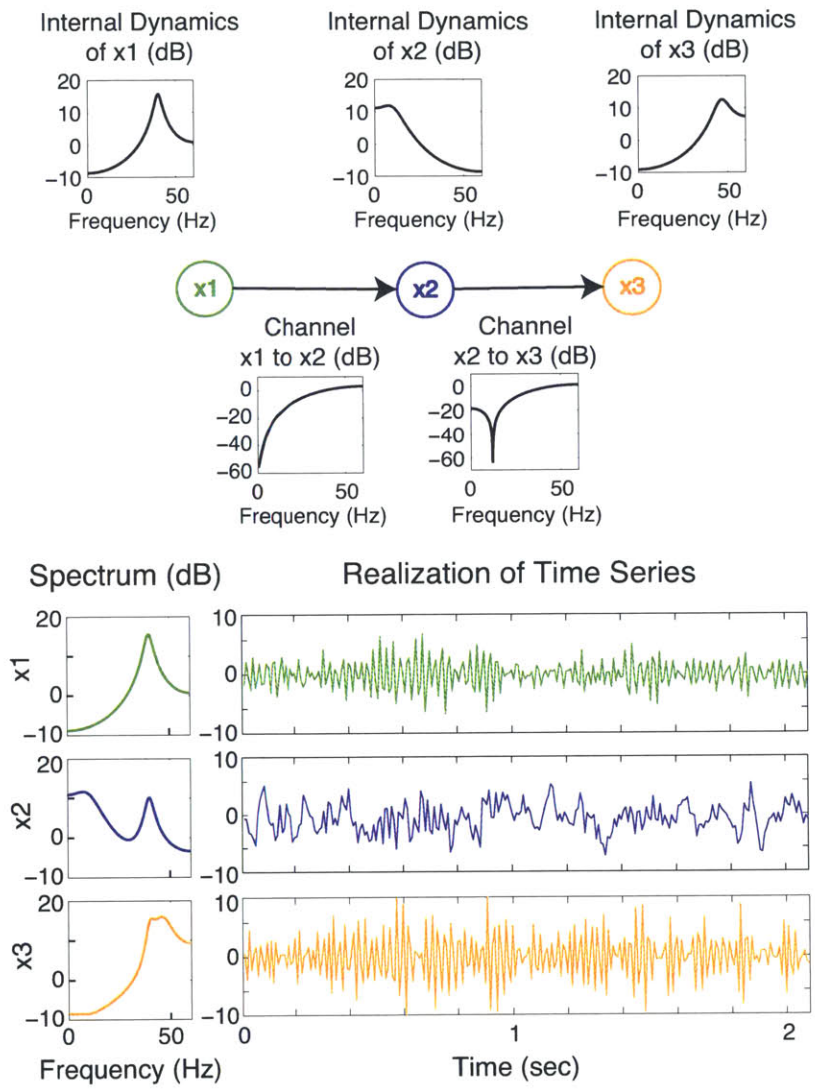

Figure 3-1: VAR(3) 3-Node Series System of Example 1. The top panel shows the network diagram of the system, with plots showing the frequency response of the internal dynamics (univariate AR behavior) of each node and the frequency response of the channels connecting the nodes. The bottom panel shows the spectrum of each node and a corresponding time series of a single realization from the system.

transmitted to node 2 with an approximate high-pass filter. Node 2, $\{x_{2,t}\}$, is driven by node 1, resonates at frequency $f_2 = 10$ Hz, and is transmitted to node 3 with an approximate high-pass filter. Node 3, $\{x_{3,t}\}$, is driven by node 2 and resonates at frequency $f_3 = 50$ Hz. All nodes also receive input from independent, white Gaussian noises, $\{w_{i,t}\}$, with zero-mean and unit-variance. The VAR equation of the system is thus

$$\begin{bmatrix} x_{1,t} \\ x_{2,t} \\ x_{3,t} \end{bmatrix} = \underbrace{\begin{bmatrix} 2r_1\cos(\theta_1) & 0 & 0 \\ -0.356 & 2r_2\cos(\theta_2) & 0 \\ 0 & -0.3098 & 2r_3\cos(\theta_3) \end{bmatrix}}_{\mathbf{A}_1} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \\ x_{3,t-1} \end{bmatrix}$$
$$+ \underbrace{\begin{bmatrix} -r_1^2 & 0 & 0 \\ 0.7136 & -r_2^2 & 0 \\ 0 & 0.5 & -r_3^2 \end{bmatrix}}_{\mathbf{A}_2} \begin{bmatrix} x_{1,t-2} \\ x_{2,t-2} \\ x_{3,t-2} \end{bmatrix}$$
$$+ \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ -0.356 & 0 & 0 \\ 0 & -0.3098 & 0 \end{bmatrix}}_{\mathbf{A}_3} \begin{bmatrix} x_{1,t-3} \\ x_{2,t-3} \\ x_{3,t-3} \end{bmatrix} + \begin{bmatrix} w_{1,t} \\ w_{2,t} \\ w_{3,t} \end{bmatrix},$$

where $r_1 = 0.9$, $r_2 = 0.7$, and $r_3 = 0.8$, $\theta_i = f_i \Delta t 2\pi$ Hz for $i = 1, 2, 3$, and the assumed sampling rate is $\frac{1}{\Delta t} = 120$ Hz.

One thousand realizations of this system were generated, each of length $N = 500$ time points (4.17 seconds). These simulated data have oscillations at the intended frequencies that are readily visible in time domain, as shown in Figure 3-1. The conditional frequency-domain causality was computed for each realization using separate full and reduced model estimates using the true model order $P = 3$, as well as model orders 6 and 20. The VAR parameters for each model order were well estimated from single realizations of the process (see Table 3.1 in Appendix 3.2.2). Model order selection was also performed using a variety of methods – including the Akaike information, Hannan-Quinn, and Schwarz criteria – all of which correctly identified,

on average, the true model order to be $P = 3$ (see Figure 3-3 in Appendix 3.2.2).

Figure 3-2 shows the distributions of causality estimates for different model orders for the truly causal connections from node 1 to node 2 and from node 2 to node 3, as well as the truly non-causal connection from node 3 to node 1. Results for the true model order 3 are shown in the first column, while those for the order 6 and 20 models are shown in the second and third columns, respectively. Each plot shows the median (blue line) and 90% central interval (blue shading) of the distribution of causality estimates. Also shown are the estimates for a single realization (red line) and the 95% significance level (black dashes) estimated by permutation [9]. The details for the significance level computation are described in Appendix 3.2.1. The fourth column shows the mean causality values for each order. The causality estimates for all cause-effect pairs for model orders 3, 6, and 20 are shown in Figures 3-4, 3-5, and 3-6, respectively, in Appendix 3.2.2.

The differences in these causality distributions reflect the trade-off in bias and variance that would be expected with increasing model order. Since the GG-causality measure is defined for infinite-order VAR models, high-order VAR models should be approximately unbiased. The similarity in the shape of the mean causality curve for high model orders (6 and 20) is consistent with this idea (Figure 3-2, column 4). A pronounced bias is evident when using the true model order 3, reflected in the large qualitative differences in the shape of the distribution compared to the model order 20 case (Figure 3-2). This is somewhat disconcerting as one might expect the causality analysis to perform best when the true model is used.

As the model order is increased, the number of parameters increases, each with their own variance, leading to increased total uncertainty in full and reduced models and in the GG-causality estimate. This has consequences for detection of causal connections, as well as for the values of the frequency-domain GG-causality estimates. For the causality detection problem, the analysis correctly identifies the significant connections under model orders 3 and 6 for all realizations. At model order 20, the variance in the causality estimates has increased, reflected in the width of the causality distribution and its null distribution, raising the 95% significance threshold

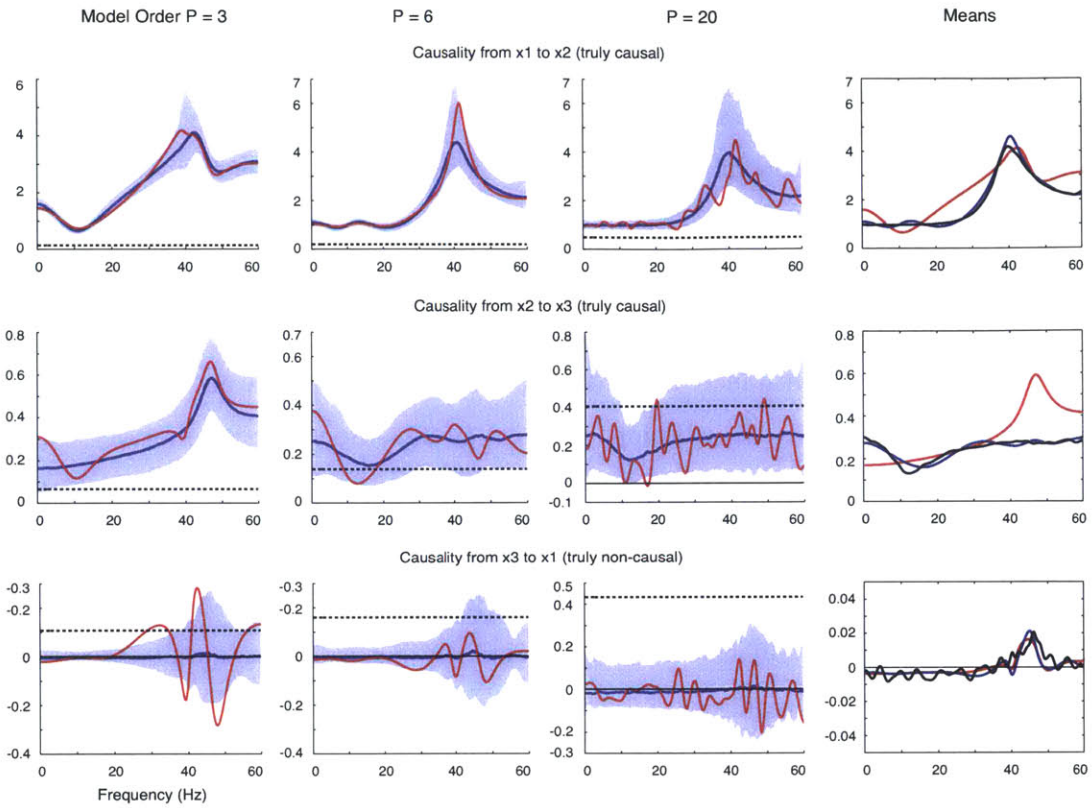

Figure 3-2: Comparison of Conditional Frequency-Domain Causality Estimates for the VAR(3) 3-Node Series System of Example 1 using Model Orders 3, 6, and 20. Rows 1, 2, and 3 show the results for the truly causal connections from node 1 to node 2 and from node 2 to node 3 and a truly non-causal connection from node 3 to node 1, respectively. Columns 1, 2, and 3 show the results from using the true model order 3 and the increased model orders 6 and 20, respectively. Each subplot shows the median (blue line) and 90% central interval (blue shading) of the distribution of causality estimates, the estimate for a single realization (red line), and the 95% significance level (black dashes) estimated by permutation. For further comparison, column 4 overlays the mean causality estimates for three model orders: 3 (red), 6 (blue), and 20 (black). The bias from using the true model order is indicated by the qualitative differences between the mean estimates of column 4 and between the distributions of column 1 compared to those of columns 2 and 3, where larger orders were used. While increasing the model order diminishes the bias, the variance is increased, indicated by the increase in the width of the distributions, the increased number of peaks and valleys of the individual realizations, and the increase in the null significance level. The additional sensitivity due to the separate full and reduced model fits is most evident from the occurrence of negative causality estimates. This occurs predominantly for the non-causal connection (row 3), but when the model order is increased to 20, instances of negative causality also occur for the truly causal connection (row2, column 3).

obtained by permutation. As a result, 7% of the realizations fail to reject the null hypothesis that no causal connection is present.

This increased variance takes a particular form, where the estimates for the individual realizations show numerous causality peaks and valleys as a function of frequency. This occurs because, as the model order increases, so too does the number of modes in the dynamical system. In frequency domain, this is reflected in the number of poles in the system, each of which represents a peak in the spectral representation of the system. Thus, with increasing model order, not only is there an increase in the variance of the causality distribution, but also an increase in the number of oscillatory components, resulting in an increase in the number of peaks and valleys in the causality estimates.

The bias in the conditional GG-causality distribution under the true model order occurs because, while GG-causality is theoretically defined using infinite histories, practical computation requires using a finite order. The reduced models are used to represent a subset of a VAR process. Unfortunately, subsets of VAR processes are not generally VAR, but instead belong to the broader class of vector auto-regressive *moving average* (VARMA) processes. VAR representations of VARMA processes are generally infinite-order. Thus, when finite-order VAR representations are used for the reduced models, as required in conditional Geweke causality computations, some terms will be omitted.

This can be seen more clearly by rewriting the full VAR model,

$$\begin{bmatrix} x_{j,t} \\ x_{i,t} \\ x_{k,t} \end{bmatrix} = \sum_{p=1}^{P} \begin{bmatrix} A_{j,j}(p) & A_{j,i}(p) & A_{j,k}(p) \\ A_{i,j}(p) & A_{i,i}(p) & A_{i,k}(p) \\ A_{k,j}(p) & A_{k,i}(p) & A_{k,k}(p) \end{bmatrix} \begin{bmatrix} x_{j,t-p} \\ x_{i,t-p} \\ x_{k,t-p} \end{bmatrix} + \begin{bmatrix} w_{j,t} \\ w_{i,t} \\ w_{k,t} \end{bmatrix}.$$

Focusing on the reduced components, the VARMA form is apparent,

$$\begin{bmatrix} x_{j,t} \\ x_{k,t} \end{bmatrix} = \underbrace{\sum_{p=1}^{P} \begin{bmatrix} A_{j,j}(p) & A_{j,k}(p) \\ A_{k,j}(p) & A_{k,k}(p) \end{bmatrix} \begin{bmatrix} x_{j,t-p} \\ x_{k,t-p} \end{bmatrix}}_{\text{Autoregressive Term}} + \underbrace{\begin{bmatrix} w_{j,t} \\ w_{k,t} \end{bmatrix} + \sum_{p=1}^{P} \begin{bmatrix} A_{j,i}(p) \\ A_{k,i}(p) \end{bmatrix} x_{i,t-p}}_{\text{Moving-Average Term}}.$$

For the reduced model, the MA terms, specifically those containing x_i, need to be appropriately projected onto the typically infinite histories of the remaining x_j and x_k components. Truncation of the reduced VAR to finite order eliminates non-zero, though diminishing, terms. Thus, when using a finite order, especially the true model order for the full system, the reduced model components can be poorly represented, leading to biased causality estimates.

The stochastic nature of the separate fits can further contribute to artifactual peaks and valleys, and even negative causality [16]. This is because when the full and reduced models are estimated separately, their frequency structure can be misaligned, producing spurious fluctuations in the resulting causality estimate. As seen in Figure 3-2, negative values are extensive for the truly non-causal connection from node 3 to node 1 and even appear at model order 20 for the truly causal connection from node 2 to node 3. This sensitivity can be quite dramatic depending on the specific system, and compounds with the increased variability due to high model order as discussed above.

This chapter has demonstrated that the practicalities of computing conditional GG-causality can lead to bias when using the true system model order and to spurious spectral peaks and valleys and weakened detection power when the order is increased. In the next chapter, a method to compute conditional GG-causality that resolves these problems is presented. Chapter 5 illustrates how, even once these issues are overcome, correct interpretation of the causality values can be problematic.

3.2 Appendix

3.2.1 Permutation Significance for Non-Causality

Significance levels for the conditional causality estimates are estimated by permutation [9]. In Example 1 above, 1000 realizations of the system were generated. For each cause-effect pair (from i to j conditional on k), 1000 permuted realizations are formed by randomly selecting realization indices I_i and I_{jk} from a uniform distribution

and recombining the I_ith realization of $\{x_{i,t}\}$ with the I_{jk}th realization of $\{x_{\{j,k\},t}\}$ to generate a permuted realization. The conditional frequency-domain causality is estimated for each permuted realization and the maximal value over all frequencies is identified. The 95% significance level over all frequencies is taken as the 95% of the 1000 maximal values. This significance level is against the null hypothesis of no difference between the causality spectrum and that of a non-causal system. It is the level below which the maximum of 95% of null causality permutations reside. This tests the whole causality waveform against the null hypothesis of no causal connection, and no assessment of significance for particular frequencies, despite the presence of peaks or valleys, can be made. The procedure can be applied to particular frequency bands of interest by identifying the 95% level within the band, but if multiple bands are tested, the appropriate corrections must be applied.

3.2.2 Additional Table and Figures

The following table and figures provide additional information about the VAR parameter and conditional frequency-domain causality estimates of the VAR(3) 3-node series system of Example 1.

Table 3.1 shows the mean and standard deviation for the VAR parameter estimates of the full model for orders 3, 6, and 20. The data show that the parameters are estimated appropriately, suggesting that the bias-variance trade-off and separate model mismatch problems described in the main text are not artifacts of poor estimation of the VAR parameters.

Figure 3-3 shows the mean values of several order selection criteria—the Akaike Information Criterion (red), the Hannan-Quinn Criterion (blue), and the Schwarz Criterion (green)—for various model orders. On average, all three criteria were minimized when using the true model order 3. This shows that for this example the true system model order would have been identified by any of the mentioned selection criteria, resulting in biased causality.

Figure 3-4 shows the causality estimates for all cause-effect pairs from using the true model order 3. Figures 3-5 and 3-6 show the estimates from using the increased

Table 3.1: Non-zero parameters for VAR(3) 3-node series system of Example 1

System Comp.	Non-Zero Param.	True Value	Order 3 mean (std.)	Order 6 mean (std.)	Order 20 mean (std.)
Node 1 Dyn.	$A_{1,1}(1)$	-0.900	-0.900 (0.044)	-0.900 (0.045)	-0.899 (0.047)
	$A_{1,1}(2)$	-0.810	-0.809 (0.053)	-0.813 (0.063)	-0.812 (0.066)
Node 2 Dyn.	$A_{2,2}(1)$	1.212	1.211 (0.044)	1.211 (0.047)	1.211 (0.050)
	$A_{2,2}(2)$	-0.490	-0.490 (0.062)	-0.490 (0.075)	-0.492 (0.079)
Node 3 Dyn.	$A_{3,3}(1)$	-1.212	-1.212 (0.042)	-1.212 (0.046)	-1.212 (0.049)
	$A_{3,3}(2)$	-0.640	-0.640 (0.059)	-0.643 (0.072)	-0.644 (0.078)
Channel 1 to 2	$A_{2,1}(1)$	-0.356	-0.355 (0.045)	-0.355 (0.046)	-0.354 (0.048)
	$A_{2,1}(2)$	0.714	0.714 (0.052)	0.714 (0.062)	0.715 (0.065)
	$A_{2,1}(3)$	-0.356	-0.355 (0.070)	-0.356 (0.084)	-0.354 (0.090)
Channel 2 to 3	$A_{3,2}(1)$	-0.310	-0.311 (0.044)	-0.310 (0.046)	-0.311 (0.049)
	$A_{3,2}(2)$	0.500	0.501 (0.062)	0.501 (0.074)	0.502 (0.078)
	$A_{3,2}(3)$	-0.310	-0.313 (0.039)	-0.313 (0.079)	-0.313 (0.085)
Noise 1	$\Sigma_{1,1}$	1	1.000 (0.067)	1.000 (0.069)	1.000 (0.073)
Noise 2	$\Sigma_{2,2}$	1	1.000 (0.064)	1.000 (0.065)	1.000 (0.069)
Noise 3	$\Sigma_{3,3}$	1	1.002 (0.063)	1.002 (0.064)	1.002 (0.068)

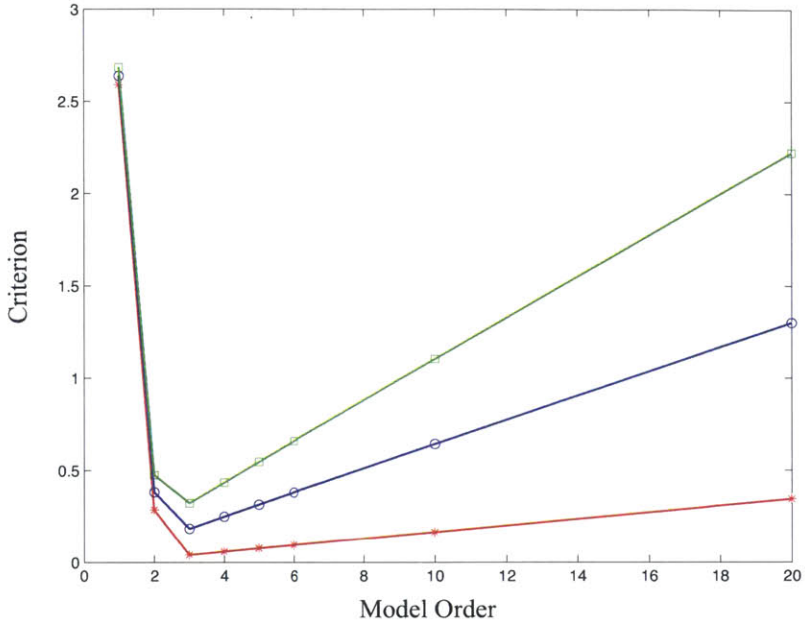

Figure 3-3: Several model order selection criteria for various model orders of VAR estimates of the VAR(3) 3-node series system of Example 1. Mean values (over 1000 realizations) of the Akaike Information Criterion (red), the Hannan-Quinn Criterion (blue), and the Schwarz Criterion (green). All criteria are minimized (on average) at the true order 3.

orders 6 and 20, respectively. Each subplot, corresponding to a cause-effect pair, shows the median (blue line) and 90% central interval (blue shading) of the distribution of the estimates, as well as the 95% significance level (black dashes) estimated by permutation and the estimate for a single realization (red line). These figures complement Figure 3-2 by showing the results for all possible connections between nodes. Comparing the figures again shows the the change in the distributions as the order is increased due to the bias using the true model order. At the same time, increasing the order increases "noisiness" of the estimates of individual realizations due to the additional spectral peaks in models.

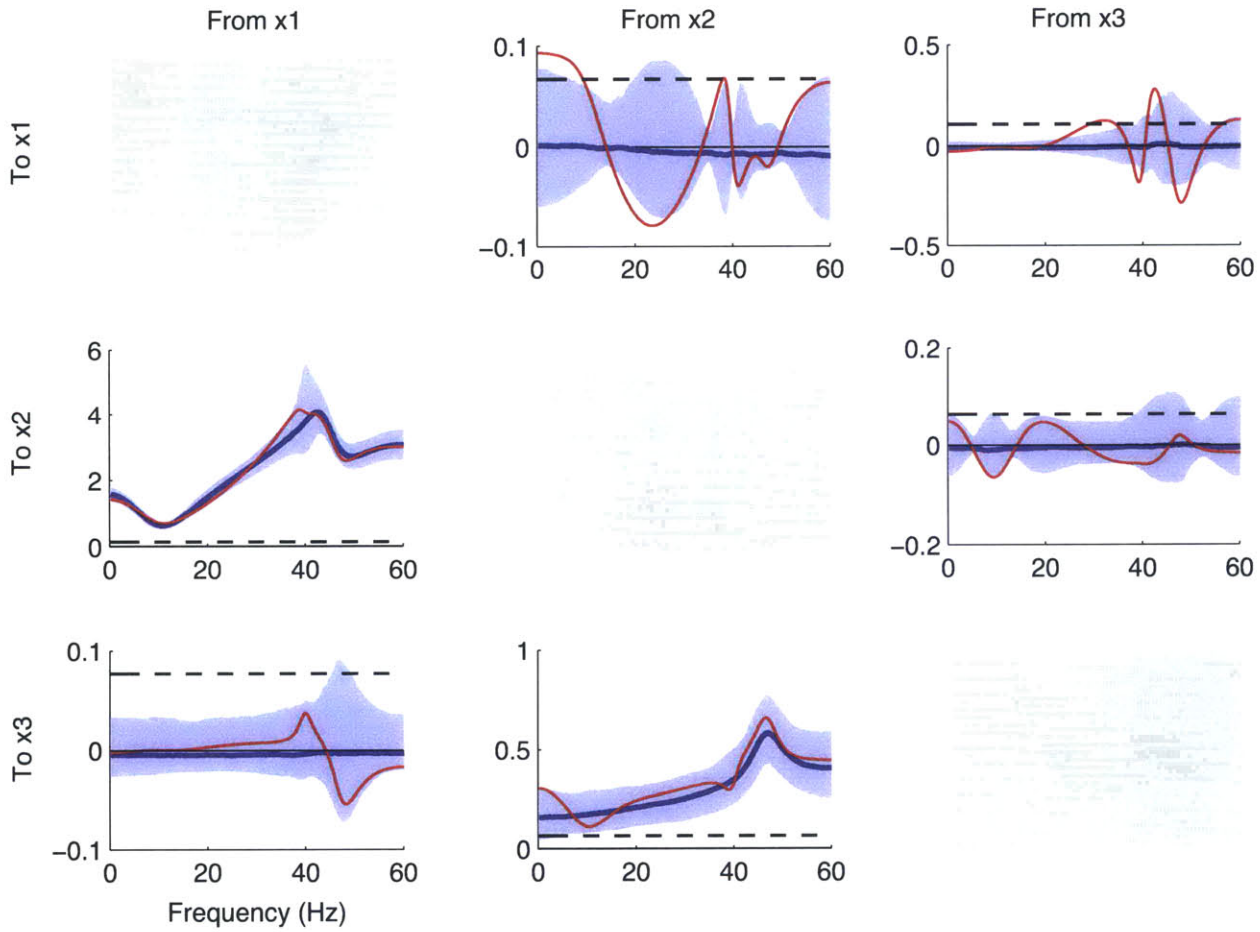

Figure 3-4: Conditional frequency-domain causality estimates for the 3-node series system from Example 1 using the true model order 3

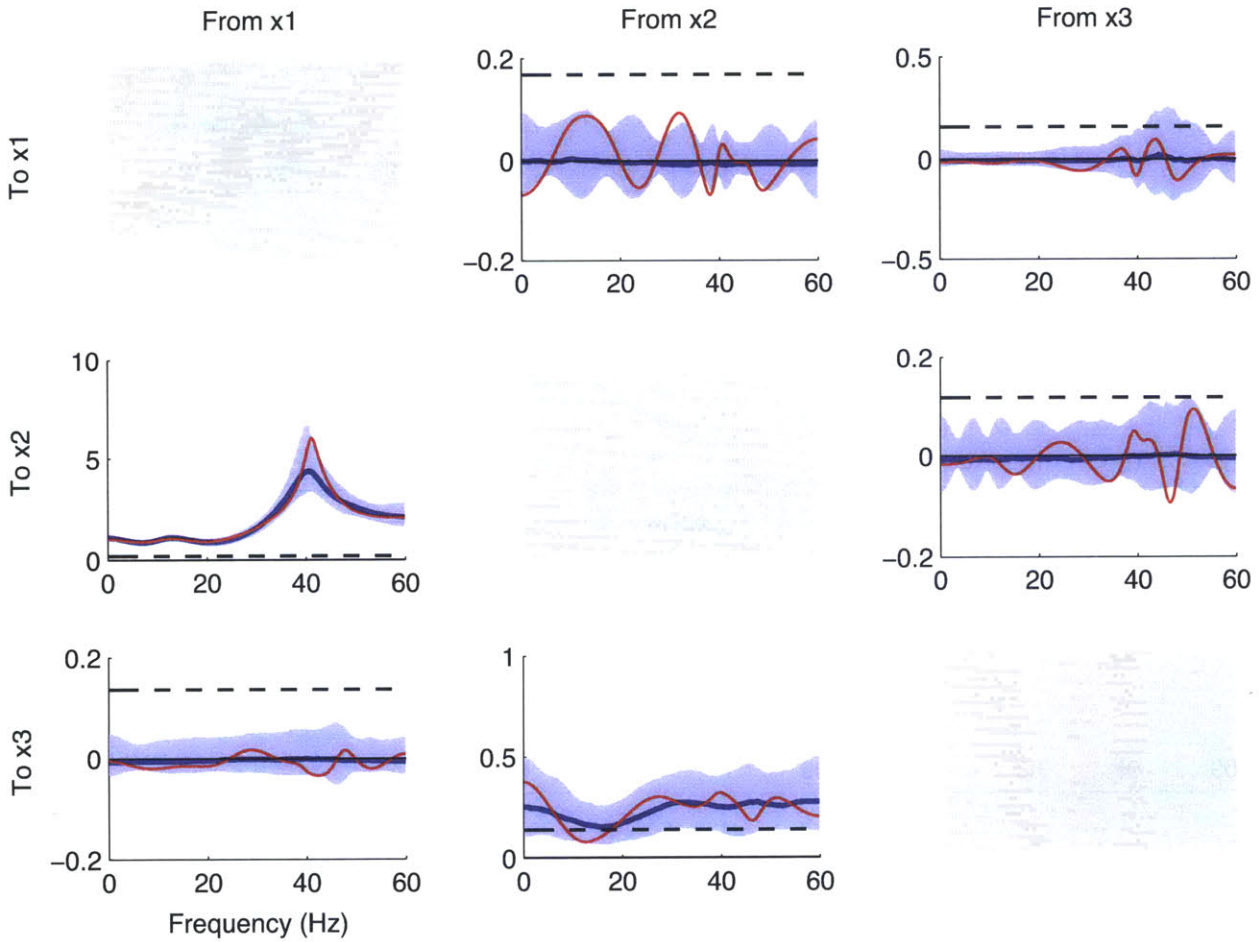

Figure 3-5: Conditional frequency-domain causality estimates for the 3-node series system from Example 1 using the increased model order 6

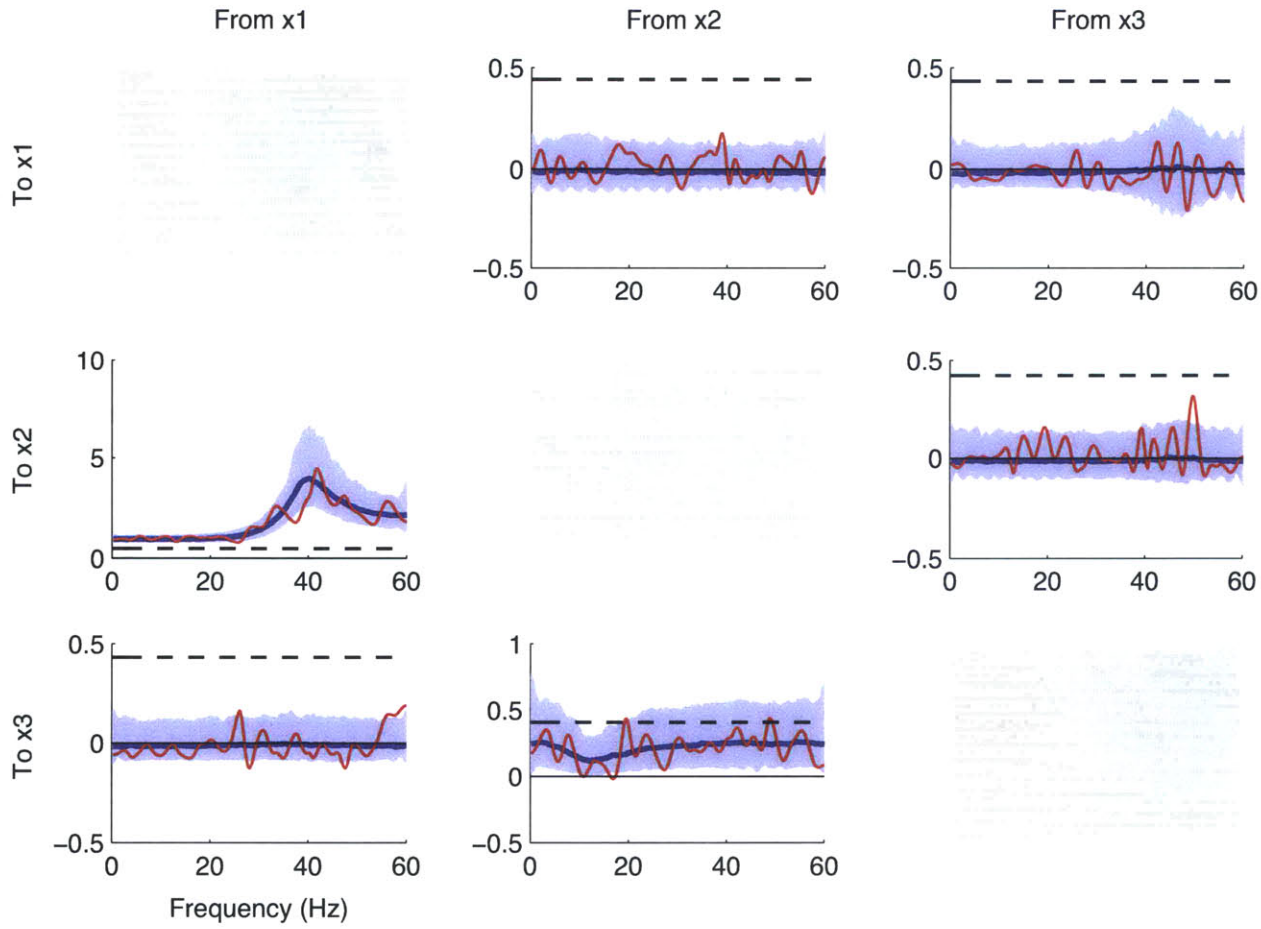

Figure 3-6: Conditional frequency-domain causality estimates for the 3-node series system from Example 1 using the increased model order 20

Chapter 4

Spectral Factorization

This chapter presents the solution to the computational problems of analyzed in the previous chapter. In order to keep the presentation self-contained, some of the background material of Chapter 2 is presented again.

4.1 Introduction

Granger causality analysis [29] is an approach to analyzing directional dependence between time series. The Geweke time- and frequency-domain measures of Granger causality [26, 27] have recently become widely applied in neuroscience. (See Chapter 1 and the discussion of Chapter 5 for applications.) Common MATLAB toolboxes [49, 7] employ methods to compute these measures. However, as detailed in the previous chapter, there are computational problems that afflict the estimation of conditional GG-causality in both the time and frequency domains.

The Granger notion of causality between time series centers on improvement of prediction. A time series $\{x_{i,t}\}$ is Granger-causal of another time series $\{x_{j,t}\}$ conditional on a background information set Ω if knowledge of the history of x_j improves the prediction of x_i over knowledge of the history of x_i and Ω alone. Typically, the background information is a set of additional time series $\{x_{k,t}\}$, and the improvement of prediction is characterized by the one-step prediction-error variances of the optimal

linear predictors, such that $\{x_{i,t}\}$ Granger-causes $\{x_{j,t}\}$ conditional on $\{x_{k,t}\}$ if

$$\text{var}\left(x_{j,t} - \hat{x}_{j,t}^{r(i)}\right) > \text{var}\left(x_{j,t} - \hat{x}_{j,t}^{f}\right),$$

where $\hat{x}_t^f = \sum_{p=1}^{\infty} A^f(p)\, x_{\{j,i,k\},t-p}$ and $\hat{x}_t^{r(i)} = \sum_{p=1}^{\infty} A^{r(i)}(p)\, x_{\{j,k\},t-p}$. Granger causality can be generalized to multi-step predictions, higher-order moments of the prediction distributions, and alternative predictors.

Geweke [26] defined unconditional measures of Granger causality in the time and frequency domains and [27] extended the definitions to include conditional time series. The time-domain Granger-Geweke (GG) causality from $\{x_{i,t}\}$ to $\{x_{j,t}\}$ conditional on $\{x_{k,t}\}$ is

$$\mathcal{F}_{x_i \longrightarrow x_j | x_k} = \ln \frac{|\text{var}\left(x_{j,t} - \hat{x}_{j,t}^{r(i)}\right)|}{|\text{var}\left(x_{j,t} - \hat{x}_{j,t}^{f}\right)|} = \ln \frac{|\Sigma_{j,j}^{r(i)}|}{|\Sigma_{j,j}^{f}|}, \tag{4.1}$$

where $|\Sigma| = \sqrt{\det \Sigma' \Sigma}$. This measure has the benefits of being non-negative and affording a spectral decomposition from which the frequency-domain measure is defined. The inclusion of conditional time series components makes it possible to distinguish between direct and indirect influences, i.e. between whether x_i directly influences x_j or whether the influence is mediated by x_k.

From the VAR representation $x_t = \sum_{p=1}^{\infty} A^f(p)\, x_{t-p} + w_t^f$, with zero-mean, temporally uncorrelated noise process $\{w_t^f\}$ of covariance Σ^f, assuming invertibility of $\left(I - \sum_{p=1}^{\infty} A^f(p)\right)$, the spectrum of x is $S_x(\lambda) = H^f(\lambda)\, \Sigma^f H^{f*}(\lambda)$, where $H^f(\lambda) = \left(I - \sum_{p=1}^{\infty} A^f(p)\, e^{-i\lambda p}\right)^{-1}$ is the full model transfer function. The frequency-domain causality measures require transforming the system by a rotation matrix to eliminate instantaneous causality, i.e. cross-covariance from Σ^f. For notational simplicity, this rotation is omitted from the presentation of the measures. They are specified in Appendix 2.4.2 and implemented in the computational studies that follow. The unconditional frequency-domain measure of causality from x_i to x_j is then defined as

$$f_{x_i \longrightarrow x_j}(\lambda) = \ln \frac{|S_{x_j}(\lambda)|}{|H_{j,j}^f(\lambda)\, \Sigma_{j,j}^f H_{j,j}^{f*}(\lambda)|} = \ln \frac{|H_{j,j}^f(\lambda)\, \Sigma_{j,j}^f H_{j,j}^{f*}(\lambda) + H_{j,i}^f(\lambda)\, \Sigma_{i,i}^f H_{j,i}^{f*}(\lambda)|}{|H_{j,j}^f(\lambda)\, \Sigma_{j,j}^f H_{j,j}^{f*}(\lambda)|}. \tag{4.2}$$

The derivation for the conditional form of frequency-domain causality relies on the time-domain equivalence demonstrated by Geweke [27], $\mathcal{F}_{x_i \longrightarrow x_j | x_k} = \mathcal{F}_{x_i w_k^{r(i)} \longrightarrow w_j^{r(i)}}$. The conditional frequency-domain causality is then defined by

$$f_{x_i \longrightarrow x_j | x_k}(\lambda) = f_{x_i w_k^{r(i)} \longrightarrow w_j^{r(i)}}(\lambda).$$

Thus, the original full model VAR must be rewritten in terms of the time series $\{w_{j,t}^{r(i)}\}$, $\{x_{i,t}\}$, and $\{w_{k,t}^{r(i)}\}$. This is done by cascading the full model with the inverse of an augmented form of the reduced model,

$$\begin{bmatrix} W_j^{r(i)}(\lambda) \\ X_i(\lambda) \\ W_k^{r(i)}(\lambda) \end{bmatrix} = \begin{bmatrix} H_{j,j}^{r(i)}(\lambda) & 0 & H_{j,k}^{r(i)}(\lambda) \\ 0 & I & 0 \\ H_{k,j}^{r(i)}(\lambda) & 0 & H_{k,k}^{r(i)}(\lambda) \end{bmatrix}^{-1} \begin{bmatrix} H_{j,j}^{f}(\lambda) & H_{j,i}^{f}(\lambda) & H_{j,k}^{f}(\lambda) \\ H_{i,j}^{f}(\lambda) & H_{i,i}^{f}(\lambda) & H_{i,k}^{f}(\lambda) \\ H_{k,j}^{f}(\lambda) & H_{k,i}^{f}(\lambda) & H_{k,k}^{f}(\lambda) \end{bmatrix} \begin{bmatrix} W_j^{f}(\lambda) \\ W_i^{f}(\lambda) \\ W_k^{f}(\lambda) \end{bmatrix} \tag{4.3}$$

$$= \begin{bmatrix} Q_{j,j}(\lambda) & Q_{j,i}(\lambda) & Q_{j,k}(\lambda) \\ Q_{i,j}(\lambda) & Q_{i,i}(\lambda) & Q_{i,k}(\lambda) \\ Q_{k,j}(\lambda) & Q_{k,i}(\lambda) & Q_{k,k}(\lambda) \end{bmatrix} \begin{bmatrix} W_j^{f}(\lambda) \\ W_i^{f}(\lambda) \\ W_k^{f}(\lambda) \end{bmatrix}. \tag{4.4}$$

Omitting the elimination of instantaneous causality for notational convenience, the spectra of $\{w_{j,t}^{r(i)}\}$ is then

$$S_{w_j^{r(i)}}(\lambda) = Q_{j,j}(\lambda)\,\Sigma_{j,j}^{f} Q_{j,j}^{*}(\lambda) + Q_{j,i}(\lambda)\,\Sigma_{i,i}^{f} Q_{j,i}^{*}(\lambda) + Q_{j,k}(\lambda)\,\Sigma_{k,k}^{f} Q_{j,k}^{*}(\lambda), \tag{4.5}$$

and the conditional frequency-domain causality is given by

$$f_{x_i \longrightarrow x_j | x_k}(\lambda) = \ln \frac{|S_{w_j^{r(i)}}(\lambda)|}{|Q_{j,j}(\lambda)\,\Sigma_{j,j}^{f} Q_{j,j}^{*}(\lambda)|} = \ln \frac{|\Sigma_{j,j}^{r(i)}|}{|Q_{j,j}(\lambda)\,\Sigma_{j,j}^{f} Q_{j,j}^{*}(\lambda)|},$$

where last equality follows from the assumed whiteness of $\{w_{j,t}^{r(i)}\}$ with covariance $\Sigma_{j,j}^{r(i)}$.

In practice, conditional GG-causality has been computed from separate full and reduced VAR model estimates of the same finite model order P, where typically the

order is chosen based on the fit of the full model according to the Akaike information criterion (AIC), or some other criteria. This is the method of the widely available MATLAB toolboxes [49, 7]. Unfortunately, this method of computation creates problems that afflict the causality estimates in both the time and frequency domains. As analyzed in Chapter 3, because subsets of VAR processes are not generally VAR, but instead VARMA, the finite VAR reduced model may not adequately represent the reduced components. This discrepancy results in biased causality estimates when the true or appropriately chosen full model order is used. The bias can be reduced by increasing the model order, but this increases the variance, introducing spurious peaks and valleys to the frequency-domain estimates, and reducing the power of tests of significant non-causality. Additionally, the stochastic nature of the separate model fits exacerbates the variance problem, leading to further peaks and valleys and even negative causality estimates.

The symptoms of this computational problem have been described previously, consisting predominantly the instances of negative causality estimates, but also the spurious peaks and valleys in the frequency domain. These problems have somewhat limited the adoption of conditional GG-causality methods compared with the more widely used unconditional methods. The realization that negative causality estimates stem from the use of separate finite model fits has been recognized previously, and has motivated other methods of computing conditional GG-causality [16, 7]. However, these methods do not utilize the correct reduced model. Because the full model represents the joint process, and the reduced model represents a subprocess, the two representations must agree in terms of correlative/spectral structure. In particular, the reduced model is derivable from the full model. The mapping between the full model parameters and the desired reduced model parameters has not been described previously.

In this chapter, a state-space form of the VAR model is used to derive the correct computation of the reduced model via spectral factorization. With this reduced model, the conditional GG-causality can be estimated without the computational problems discussed in Chapter 3, since only a single full model is estimated. This

method is demonstrated using a simulated example, showing that it eliminates negative GG-causality estimates, and that using the true model order results in estimates of GG-causality that are unbiased with minimal variance. Thus, this approach addresses the fundamental problems associated with GG-causality that were analyzed in Chapter 3.

4.2 Materials and Methods

4.2.1 Standard Computation of Conditional GG-Causality

The standard computation of conditional GG-causality uses separate estimates of a full VAR model, which includes all time series components, and a reduced VAR model, from which the putatively causal component time series is omitted. In the case of analyzing the causality from $\{x_{i,t}\}$ to $\{x_{j,t}\}$ conditional on $\{x_{k,t}\}$, the full and reduced models of order P are

$$\begin{bmatrix} x_{j,t} \\ x_{i,t} \\ x_{k,t} \end{bmatrix} = \sum_{p=1}^{P} \begin{bmatrix} A^f_{j,j}(p) & A^f_{j,i}(p) & A^f_{j,k}(p) \\ A^f_{i,j}(p) & A^f_{i,i}(p) & A^f_{i,k}(p) \\ A^f_{k,j}(p) & A^f_{k,i}(p) & A^f_{k,k}(p) \end{bmatrix} \begin{bmatrix} x_{j,t-p} \\ x_{i,t-p} \\ x_{k,t-p} \end{bmatrix} + \begin{bmatrix} w^f_{j,t} \\ w^f_{i,t} \\ w^f_{k,t} \end{bmatrix}, \tag{4.6}$$

and

$$\begin{bmatrix} x_{j,t} \\ x_{k,t} \end{bmatrix} = \sum_{p=1}^{P} \begin{bmatrix} A^{r(i)}_{j,j}(p) & A^{r(i)}_{j,k}(p) \\ A^{r(i)}_{k,j}(p) & A^{r(i)}_{k,k}(p) \end{bmatrix} \begin{bmatrix} x_{j,t-p} \\ x_{k,t-p} \end{bmatrix} + \begin{bmatrix} w^{r(i)}_{j,t} \\ w^{r(i)}_{k,t} \end{bmatrix}, \tag{4.7}$$

respectively, where $\{w^f_t\}$ and $\{w^{r(i)}_t\}$ are white noise processes with covariance matrices Σ^f with covariance $\Sigma^{r(i)}$.

4.2.2 Computation using Reduced Model from Spectral Factorization

In the assessment of Granger causality and the computation of the Geweke measures, the intent is to compare the uncertainty of the optimal predictor of the reduced process given its own complete history to the uncertainty of the optimal predictor when the history of the putatively causal time series is included, i.e. to compare the reduced model, $x_t^{r(i)} = \sum_{p=0}^{\infty} h_p^{r(i)} w_{t-p}^{r(i)}$, to the full model $x_t^f = \sum_{p=0}^{\infty} h_p^f w_{t-p}^f$. These two models must agree on the represented correlation structure and spectra of the reduced process, and given a VAR(P) full model of the joint process, the VARMA structure of the reduced process is determined. The difficulty has been determining the corresponding VAR reduced model of the same dimension as the reduced process. The easiest solution to obtain the desired unique reduced model that is to compute it numerically by spectral factorization of the full model in state-space form.

Treat the full VAR(P) model as an underlying state process and write it in VAR(1) form,

$$\begin{aligned}\overline{x}_{t+1} =& F\overline{x}_t + Gw_t^f \\ =& \begin{bmatrix} A^f(1) & \cdots & A^f(2) & A^f(P) \\ & \mathrm{I} & & 0 \end{bmatrix} \begin{bmatrix} x_t \\ \vdots \\ x_{t-(P-1)} \end{bmatrix} + \begin{bmatrix} \mathrm{I} \\ 0 \end{bmatrix} w_t^f,\end{aligned}$$

where $\overline{x}_t = \begin{bmatrix} x_t^T & x_{t-1}^T & \cdots & x_{t-(P-1)}^T \end{bmatrix}^T$ is the augmented state of lagged values. In the case of the full model, the state is directly observed, $y_t^f = x_t$, so the observation noise is zero, $v_t^f = 0$:

$$\begin{aligned}y_t^f =& U^f\overline{x}_t + v_t^f \\ =& \begin{bmatrix} \mathrm{I} & 0 & \ldots & 0 \end{bmatrix} \begin{bmatrix} x_t \\ \vdots \\ x_{t-(P-1)} \end{bmatrix} = x_t,\end{aligned}$$

where U^f is a selection matrix that selects the top block of the state. The spectrum of the full process $\{y_t^f\}$ from this state-space form is equal to the spectrum of the full model VAR(P) $\{x_t\}$,

$$\begin{aligned} S_{y^f}(z) =& U^f (z\mathrm{I} - F)^{-1} G\Sigma^f G^* \left(z^{-1}\mathrm{I} - F^*\right)^{-1} U^{f*} && (4.8)\\ =& H^f(z)\,\Sigma^f H^{f*}(z) && (4.9)\\ =& S_x(z), && (4.10) \end{aligned}$$

where $H^f(z) = \left(I - \sum_{p=1}^{P} A^f(p)\, z^{-p}\right)^{-1}$.

In the case of the reduced model, only a subset the state components are observed, $y_t^{r(i)} = \begin{bmatrix} x_{j,t}^T & x_{k,t}^T \end{bmatrix}^T$, but they are still observed without noise, $v_t^{r(i)} = 0$:

$$\begin{aligned} y_t^{r(i)} =& U^{r(i)}\overline{x}_t + v_t^{r(i)} \\ =& \begin{bmatrix} \begin{bmatrix} \mathrm{I}_{|j|} & 0_{|k|} & \mathrm{I}_{|k|} \end{bmatrix} & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} x_t \\ \vdots \\ x_{t-(P-1)} \end{bmatrix} = \begin{bmatrix} x_{j,t} \\ x_{k,t} \end{bmatrix}, \end{aligned}$$

where $U^{r(i)}$ is a selection matrix that selects the observed components from the top block of the state. The spectrum of the reduced process $\{y_t^{r(i)}\}$ is

$$S_{y^{r(i)}}(z) = U^{r(i)} (z\mathrm{I} - F)^{-1} G\Sigma^f G^* \left(z^{-1}\mathrm{I} - F^*\right)^{-1} U^{r(i)*}. \tag{4.11}$$

In seeking the reduced VAR model, one is looking for a causal, stable, invertible system driven by a noise process of the same dimension as the observations, i.e.

$$S_{y^{r(i)}}(z) = H^{r(i)}(z)\,\Sigma^{r(i)} H^{r(i)*}\left(z^{-*}\right). \tag{4.12}$$

This spectral factorization is well known [33, 1] and given by

$$H^{r(i)}(z) = U^{r(i)} (z\mathrm{I} - F)^{-1} K + \mathrm{I}, \tag{4.13}$$

where

$$\Sigma^{r(i)} = U^{r(i)}\overline{\Sigma}^{r(i)}U^{r(i)*} \tag{4.14}$$

is the prediction-error covariance of the model,

$$K = \left(F\overline{\Sigma}^{r(i)}U^{r(i)*}\right)\left(\Sigma^{r(i)}\right)^{-1} \tag{4.15}$$

is the Kalman gain, and $\overline{\Sigma}^{r(i)}$ is the solution to the Discrete-Time Algebraic Riccati Equation (DARE),

$$\overline{\Sigma}^{r(i)} = F\overline{\Sigma}^{r(i)}F^* + G\Sigma^f G^* - K\Sigma^{r(i)}K^*. \tag{4.16}$$

This spectral factorization gives the reduced model in the following innovations form

$$y_t = U^{r(i)}\overline{x}_t + w_t^{r(i)} \tag{4.17}$$

$$\overline{x}_{t+1} = F\overline{x}_t + Kw_t^{r(i)}. \tag{4.18}$$

The frequency-domain conditional GG-causality is computed using the reduced model $H^{r(i)}(z)$ from Eqn. 4.13 in Eqn. 4.3. The time-domain conditional GG-causality can be computed either by integrating the frequency-domain or directly by obtaining the DARE solution $\overline{\Sigma}^{r(i)}$ and using the prediction-error covariance, $\Sigma^{r(i)} = U^{r(i)}\overline{\Sigma}^{r(i)}U^{r(i)*}$, in Eqn. 4.1.

Note that the specific form of the above state-space model, that of a partially observed VAR(P) process, eliminates terms from the general spectral factorization equation, notably those involving the observation noise covariance matrix and the observation-state noise cross-covariance matrix. Thus, the notation used in the above presentation is suggestive of the GG-causality computation. Though the observation noise covariance is zero, the spectra exist (being those of a VAR process), all inverses involved exist, and the conditions for solution are satisfied. The spectral factorization equations for the general state-space model and conditions for its solution are included in Appendix 4.5.1.

Example 2: VAR(3) 3-Node Series System

This example analyzes time series generated by the VAR(3) system shown in Figure 4-1. In this system, node 1, $\{x_{1,t}\}$, resonates at frequency $f_1 = 40$ Hz and is transmitted to node 2 with an approximate high-pass filter. Node 2, $\{x_{2,t}\}$, is driven by node 1, resonates at frequency $f_2 = 10$ Hz, and is transmitted to node 3 with an approximate high-pass filter. Node 3, $\{x_{3,t}\}$, is driven by node 2 and resonates at frequency $f_3 = 50$ Hz. All nodes also receive input from independent, white Gaussian noises, $\{w_{i,t}\}$, with zero-mean and unit-variance. The VAR equation of the system is thus

$$\begin{aligned}
\begin{bmatrix} x_{1,t} \\ x_{2,t} \\ x_{3,t} \end{bmatrix} = & \underbrace{\begin{bmatrix} 2r_1\cos(\theta_1) & 0 & 0 \\ -0.356 & 2r_2\cos(\theta_2) & 0 \\ 0 & -0.3098 & 2r_3\cos(\theta_3) \end{bmatrix}}_{\mathbf{A}_1} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \\ x_{3,t-1} \end{bmatrix} \\
& + \underbrace{\begin{bmatrix} -r_1^2 & 0 & 0 \\ 0.7136 & -r_2^2 & 0 \\ 0 & 0.5 & -r_3^2 \end{bmatrix}}_{\mathbf{A}_2} \begin{bmatrix} x_{1,t-2} \\ x_{2,t-2} \\ x_{3,t-2} \end{bmatrix} \\
& + \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ -0.356 & 0 & 0 \\ 0 & -0.3098 & 0 \end{bmatrix}}_{\mathbf{A}_3} \begin{bmatrix} x_{1,t-3} \\ x_{2,t-3} \\ x_{3,t-3} \end{bmatrix} + \begin{bmatrix} w_{1,t} \\ w_{2,t} \\ w_{3,t} \end{bmatrix},
\end{aligned}$$

where $r_1 = 0.9$, $r_2 = 0.7$, and $r_3 = 0.8$, $\theta_i = f_i \Delta t 2\pi$ Hz for $i = 1, 2, 3$, and the assumed sampling rate is $\frac{1}{\Delta t} = 120$ Hz.

One thousand realizations of this system were generated, each of length $N = 500$ time points (4.17 seconds). These simulated data have oscillations at the intended frequencies that are readily visible in time domain, as shown in Figure 4-1. The conditional frequency-domain causality was computed for each realization using both methods—separate model fits and spectral factorization—using the true model order $P = 3$, as well as model orders 6, 10, and 20.

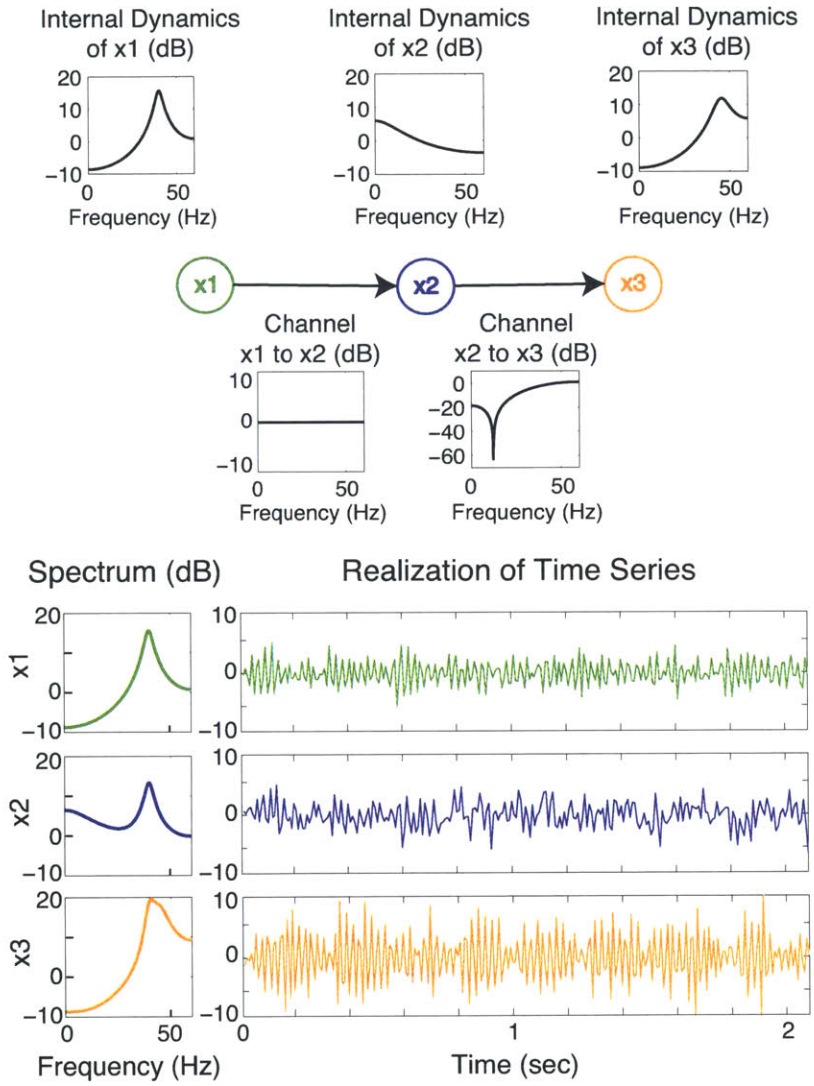

Figure 4-1: Example 2: VAR(3) 3-Node Series System. The top panel shows the network diagram of the system, with subplots showing the internal dynamics (univariate AR behavior) of each node and the behavior of the channels connecting the nodes. The bottom panel shows the spectrum of each node and a single realization of the corresponding time series.

4.3 Results

Figures 4-2, 4-3, and 4-4 compare the distributions of causality estimates obtained from the standard computation using separate model fits to those obtained using spectral factorization for the truly causal connections from node 1 to node 2 and from node 2 to node 3 and a truly non-causal connection from node 3 to node 1, respectively. Estimates obtained using separate full and reduced model fits are shown in the left columns of each figure, and estimates from using spectral factorization are shown in the right columns. Results obtained using models of the true order 3 are shown in the first row, while those for the order 6, 10, and 20 models are shown in the second, third, an fourth rows, respectively. Each plot shows the median (blue line), the 90% central interval (blue shading), and the minimum (black dashes) of the distribution of causality estimates. In addition, estimates for a single realization (red line) are shown. The causality estimates for the remaining cause-effect pairs (from node 1 to node 3, from node 2 to node 1, and from node 3 to node 2) are shown in Figures 4-5, 4-6, and 4-7, respectively, in Appendix 4.5.2.

As seen in the figures, the computational problems that affect the separate model fit estimation of GG-causality, analyzed in Chapter 3 and seen here again in the left columns, do not affect the estimation via the spectral factorization in the right columns. Most noticeably, the negative values[1] that occur for the causality estimates of each cause-effect pair for every realization when using the separate model fits do not occur for the estimates using spectral factorization. More importantly, whereas the separate model fit method faces a peculiar bias-variance trade-off, such that using the true full model order results in biased causality estimates, but increasing the model order increases the variance and introduces spurious peaks and valleys in the individual estimates (compare top left subplots to bottom left subplots), using the true model order with the spectral factorization method produces unbiased estimates with the minimal variance (top right subplots).

[1]Actually, for model order 20, several realizations did result in negative values, as the dotted curves dip slightly below zero in the bottom right subplots of Figures 4-3, 4-4, 4-5, 4-6, and 4-7. However, this is most likely a numerical issue due to excessively large model order.

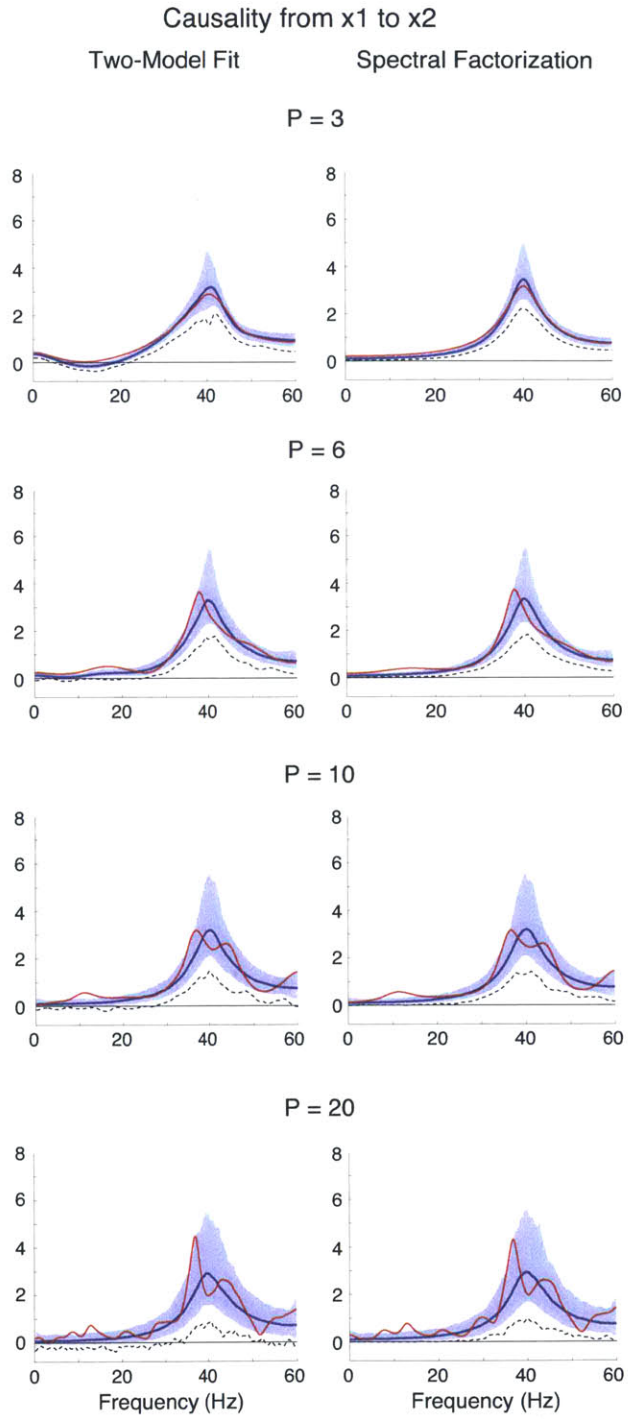

Figure 4-2: Example 2, Causality estimates from node 1 to node 2.

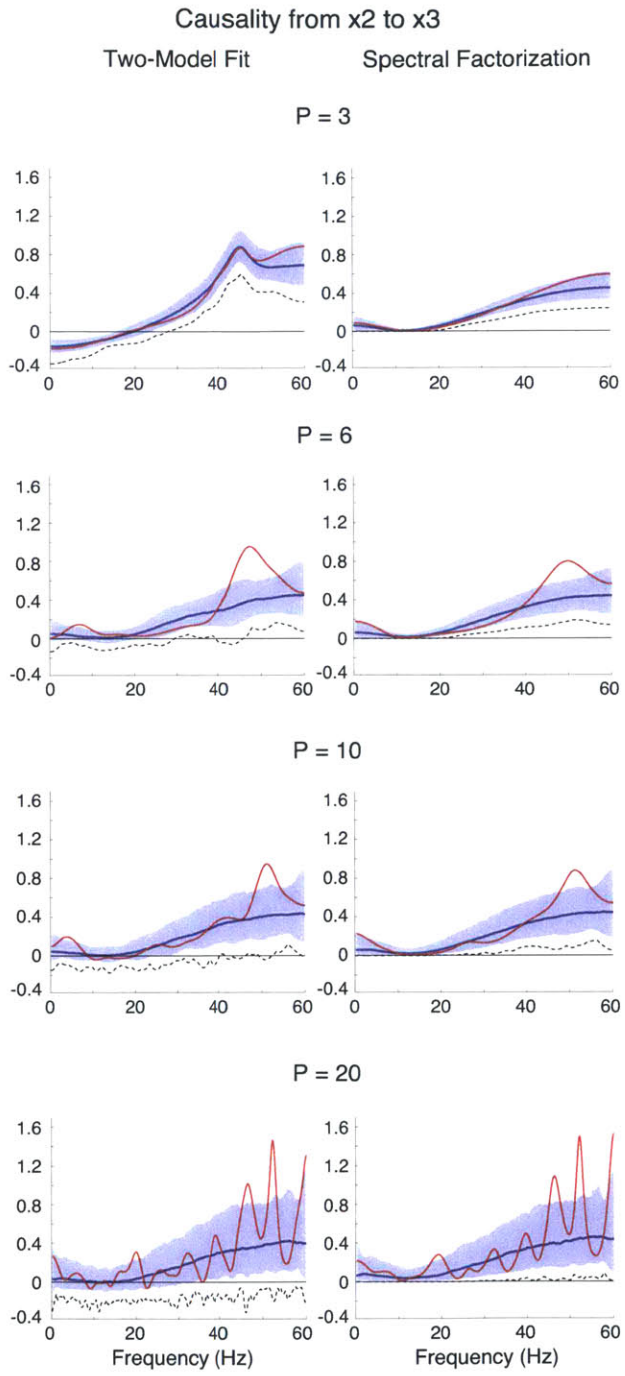

Figure 4-3: Example 2, Causality estimates from node 2 to node 3.

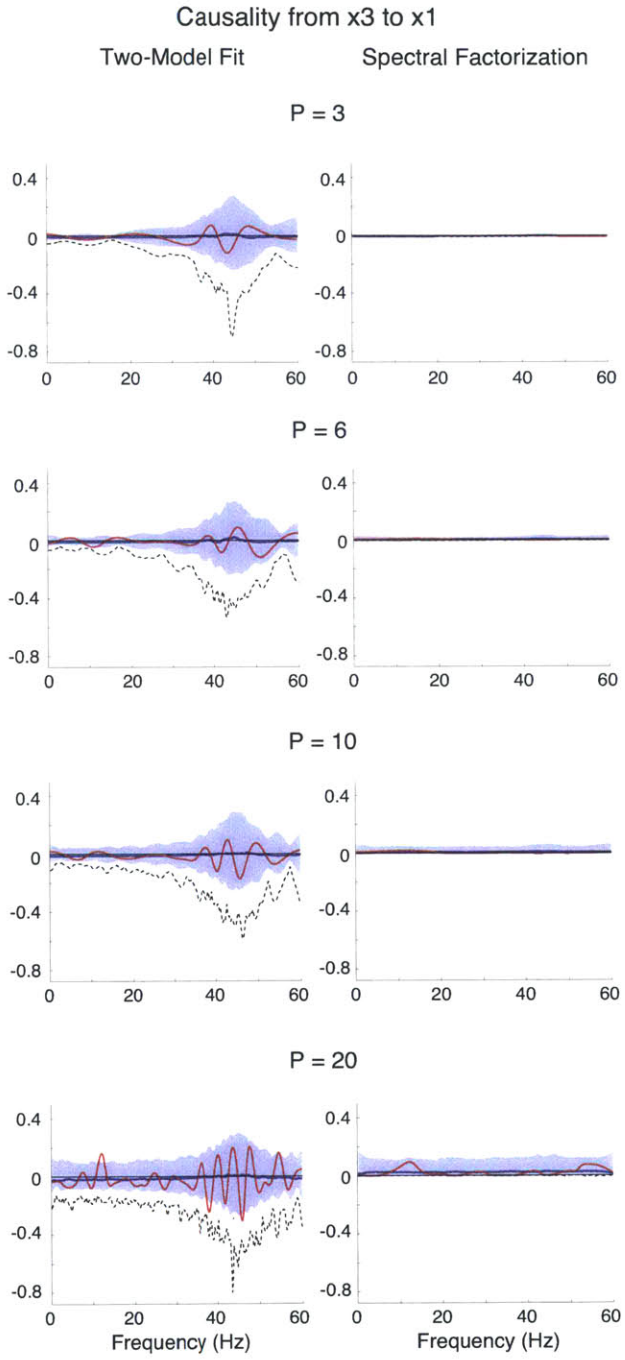

Figure 4-4: Example 2, Causality estimates from node 3 to node 1.

4.4 Discussion

The method presented above is thus the correct computation of the intended reduced model and subsequent GG-causality estimates. Using the spectral factorization of the state-space form of the full model to compute the corresponding reduced model allows estimation of GG-causality that avoids the computational problems analyzed in Chapter 3. With this method, one can estimate the GG-causality using the true (or otherwise appropriate) model order without bias, with minimal variance, and without the spurious peaks and valleys and negative values encountered with separate model fits.

Two alternative methods to the separate model fit computation have been proposed previously. Barnett et al. [7] proposed fitting the reduced model and then using the correlation and cross-correlation of the omitted time series to directly compute the combined transfer function $Q(\lambda)$ in Eqn. 4.4. However, this method still faces the problem of bias-variance trade-off. Barnett et al. [7] recommend using sufficiently large model order to represent the reduced process, but this will still lead to the increased variance, spurious peaks and valleys of frequency-domain estimates, and reduced detection power. Chen et al. [16] proposed forming a VARMA reduced model from the estimated full model parameters. The step of removing instantaneous causality is then applied to each frequency of the colored noise of the reduced model. However, this reduced model is somewhat ad hoc and not the reduced model required for evaluation of GG-causality.

This spectral factorization approach resolves the computational problems analyzed in Chapter 3. However, the interpretation of GG-causality must still be considered carefully. This question of interpretation is analyzed in Chapter 5.

4.5 Appendix

4.5.1 Spectral Factorization of General State-Space Models

This section shows conditions and equations for canonical spectral factorization of a general state-space model. These details are from [33, Chapter 8], specifically Theorems 8.3.1 and 8.3.2. Other details, such as proofs and derivations, can also be found there. See also [1].

For the general state-space model,

$$x_{t+1} = Fx_t + Gw_t \tag{4.19}$$

$$y_t = Hx_t + v_t \tag{4.20}$$

where $E\begin{bmatrix} w_t^T & v_t^T \end{bmatrix}^T = 0$ and

$$E\begin{bmatrix} w_{t_1} \\ v_{t_1} \end{bmatrix}\begin{bmatrix} w_{t_2} & v_{t_2} \end{bmatrix} = \begin{bmatrix} GQG^* & GS \\ S^*G^* & R \end{bmatrix}\delta_{t_1-t_2}, \tag{4.21}$$

if F is stable, the system is controllable on the unit circle,

$$\begin{bmatrix} GQG^* & GS \\ S^*G^* & R \end{bmatrix} \geq 0, \quad \text{and} \quad R > 0, \tag{4.22}$$

then the DARE equation

$$P = FPF^* + GQG^* - K_pR_eK_p^*, \tag{4.23}$$

where

$$K_p = (FPH^* + GS)\,R_e^{-1} \quad \text{and} \quad R_e = R + HPH^*, \tag{4.24}$$

has unique solution P such that $F - K_pH$ is stable. This stabilizing solution is positive-semi-definite such that R_e is positive-definite.

For a system satisfying the above conditions, with a spectrum of the form

$$S_y(z) = \begin{bmatrix} H(zI-F)^{-1} & I \end{bmatrix} \begin{bmatrix} GQG^* & GS \\ S^*G^* & R \end{bmatrix} \begin{bmatrix} (z^{-1}I-F^*)^{-1}H^* \\ I \end{bmatrix}, \tag{4.25}$$

the canonical spectral factorization is

$$S_y(z) = L(z) R_e L^*\left(z^{-*}\right), \tag{4.26}$$

such that $L(\infty) = I$ and $R_e > 0$, where

$$L(z) = I + H(zI-F)^{-1}K_p \quad \text{and} \quad L^{-1}(z) = I - H(zI-F+K_pH)^{-1}K_p. \tag{4.27}$$

P is the unique positive-semi-definite DARE solution, and because both F and $F - K_pH$ are stable, $L(z)$ is minimum-phase.

4.5.2 Additional Figures

Figures 4-5, 4-6, and 4-7 show the causality estimates for the remaining truly non-causal cause-effect pairs (from node 1 to node 3, from node 2 to node 1, and from node 3 to node 2, respectively) for Example 2. Each figure compares the distributions of causality estimates obtained from the standard computation using separate model fits to those obtained from the proposed computation using spectral factorization. Estimates from using separate full and reduced model fits are shown in the left columns of each figure, and estimates from using spectral factorization are shown in the right columns. Results from using models of the true order 3 are shown in the first row, while those for the order 6, 10, and 20 models are shown in the second, third, an fourth rows, respectively. Each plot shows the median (blue line), the 90% central interval (blue shading), and the minimum (black dashes) of the distribution of causality estimates. In addition, estimates for a single realization (red line) are shown.

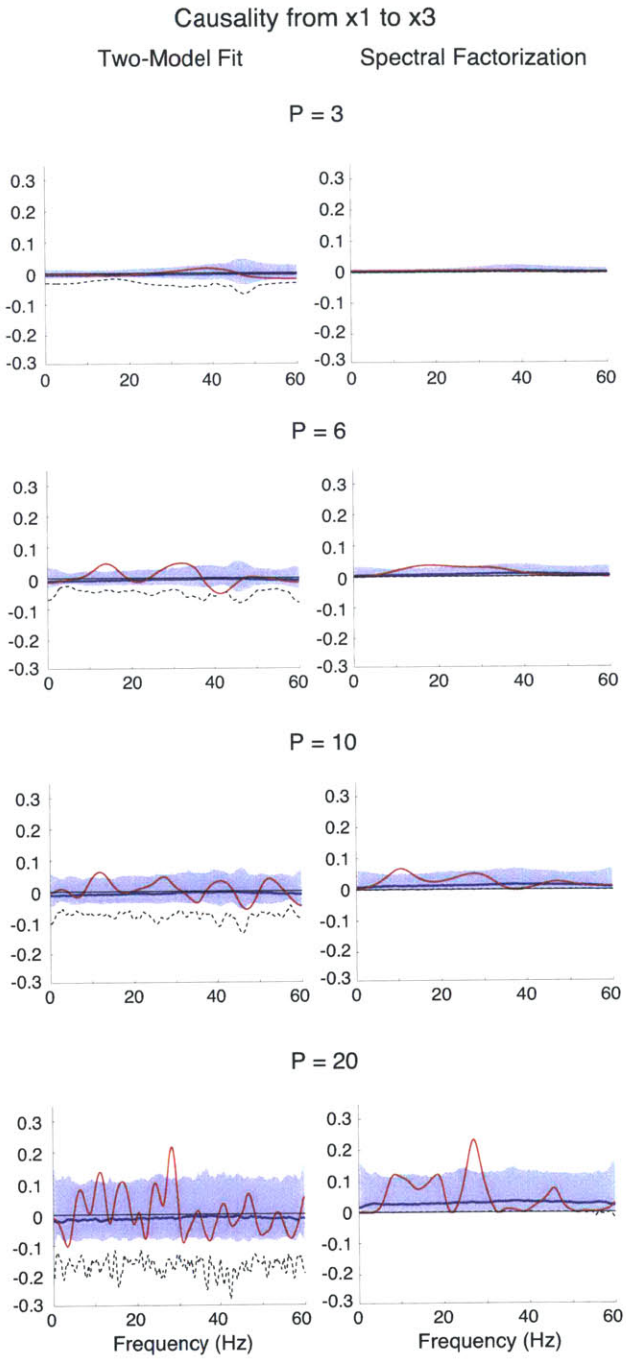

Figure 4-5: Example 2, Causality estimates from node 1 to node 3.

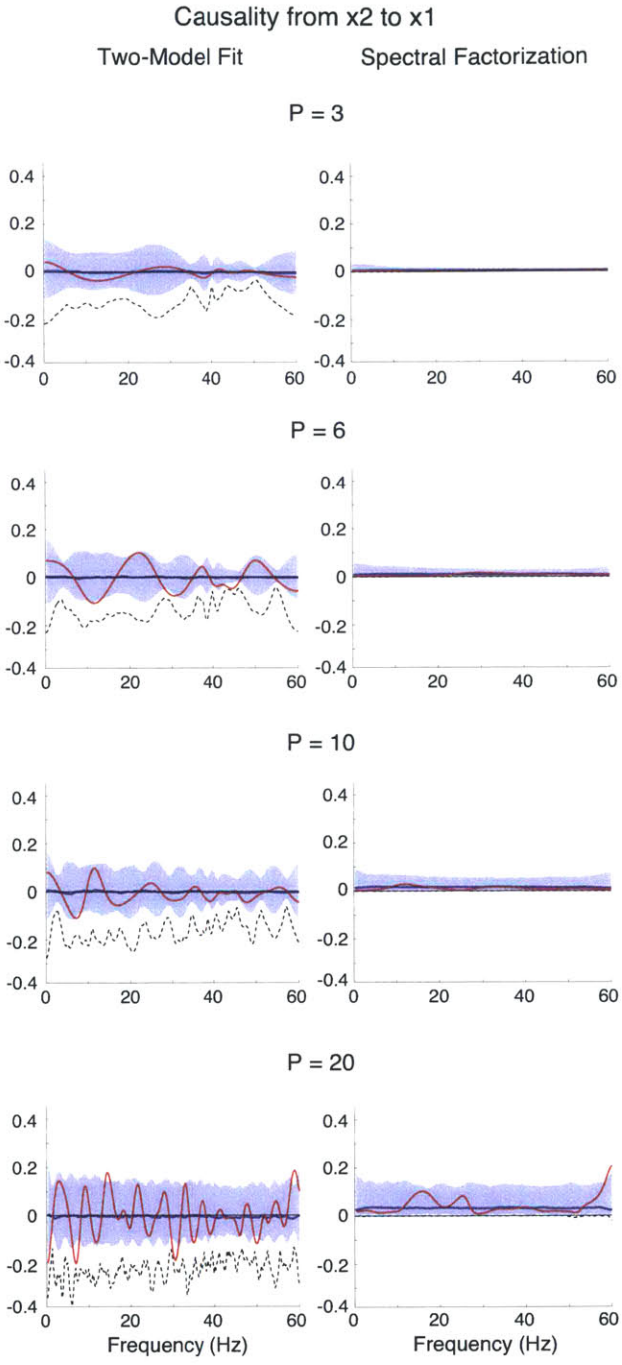

Figure 4-6: Example 2, Causality estimates from node 2 to node 1.

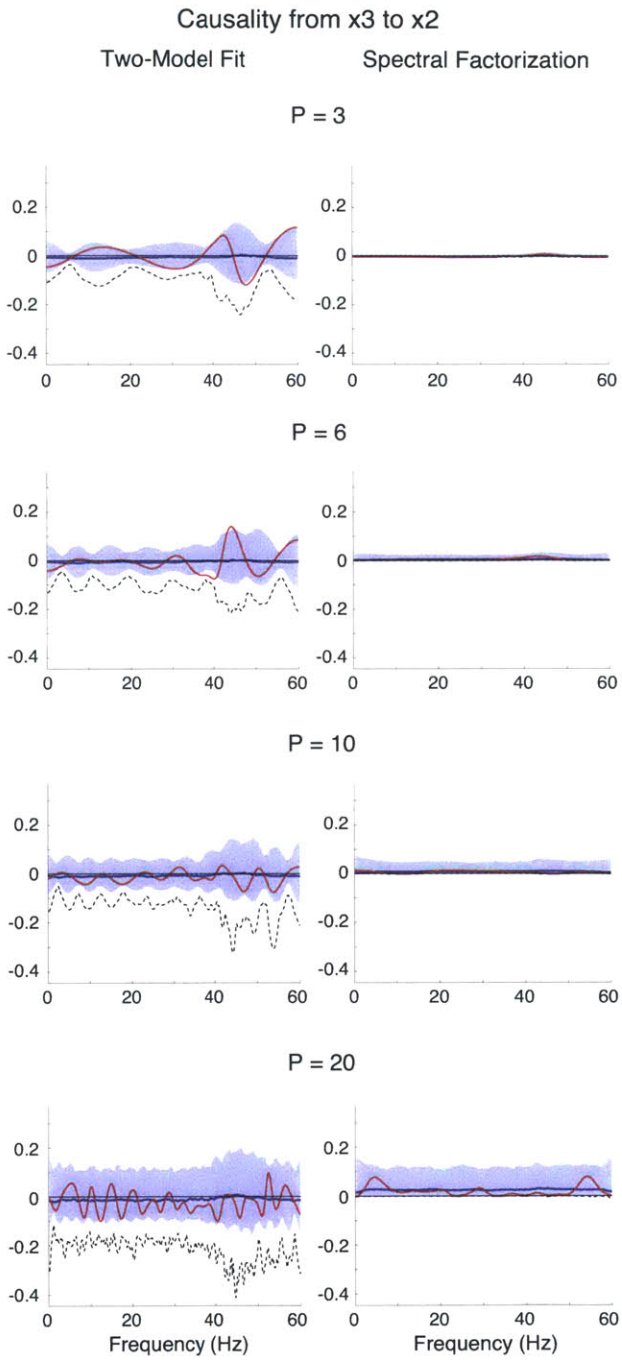

Figure 4-7: Example 2, Causality estimates from node 3 to node 2.

Chapter 5

Interpretational Problems

5.1 How are the Functional Properties of the System Reflected in GG-Causality?

Given a frequency-domain causality spectrum over a range of frequencies, a face value interpretation would be that peaks in the waveform represent frequency bands that are more causal or effective, while valleys represent frequency bands of lesser effect. Similarly, for the time-domain causality, larger values would indicate stronger effect than smaller values both within a system and in comparison to similar systems. Given this interpretation, a number of questions arise. When analyzing a time series dataset with a causality measure, what do the causality values say about the behavior of the generative system? And when comparing times series datasets from different systems, what do the causality values say about the relationship between the generative systems?

The causality measure, as a function of the system parameters, is a combination of the component dynamics of the system, merging the separate aspects of system behavior into a single value or a single-valued function of frequency. In this chapter, the relationship between system behavior and causality values is examined by varying the parameters of different parts of the system—transmitters, receivers, and channels—and assessing the dependence of the causality on the component behaviors.

It will be shown that GG-causality reflects only the dynamics of the transmitter node and channel, with no dependence on the dynamics of the receiver node. Further, the overall scaling of the GG-causality measure, as determined by the specific values of the estimated VAR parameters, can complicate comparisons of functional system properties within and between systems. These factors suggest that, even in simple bivariate AR systems, GG-causality may be prone to misinterpretation, and may not reflect the intuitive notions of causality most often associated with these methods.

Example 3: Transmitter-Receiver Pair with Varying Resonance Alignments

This example investigates the behavior of the Geweke causality measure as a function of the system parameters. It examines how the causality values change as the system's frequency structure varies, to determine if the changes in the causality values agree with the face value interpretation of causality described in Chapter 1. The analysis is structured to avoid the computational problems described in Chapter 3. Specifically, a bivariate case is analyzed, such that the unconditional causality can be computed correctly from a single model using the true VAR parameters. By using the true VAR parameters, the uncertainties associated with VAR parameter estimation are also eliminated. Thus, the sole focus in this example is the functional relationship between system structure and the causality measure.

The causality values of a set of similar unidirectional two-node systems, all with the structure shown in the left column of Figure 5-1, are compared. To isolate the influence of transmitter and receiver behaviors, the channel is fixed as all-pass (i.e., all frequencies pass through the channel without influencing the signal amplitude). In these systems, node 1, $\{x_{1,t}\}$, is driven by a white noise input, $\{w_{1,t}\}$, resonates at frequency f_1, and is transmitted to node 2. Node 2, $\{x_{2,t}\}$, is driven by both node 1 and a white noise input, $\{w_{2,t}\}$, and resonates at frequency f_2. The channel is fixed as all-pass unit delay, but the transmitter and receiver resonance frequencies are varied. The set of transmitter resonance frequencies is $f_1 \in \{10, 20, 30, 40, 50\}$ Hz. The set of receiver resonance frequencies is $f_2 \in \{10, 30, 50\}$ Hz. The VAR equation of these

systems is thus

$$\begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} = \begin{bmatrix} 2r_1 \cos(\theta_1) & 0 \\ 1 & 2r_2 \cos(\theta_2) \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} + \begin{bmatrix} -r_1^2 & 0 \\ 0 & -r_2^2 \end{bmatrix} \begin{bmatrix} x_{1,t-2} \\ x_{2,t-2} \end{bmatrix} + \begin{bmatrix} w_{1,t} \\ w_{2,t} \end{bmatrix},$$

where $r_i = 0.65$ and $\theta_i = f_i \Delta t 2\pi$ Hz for $i = 1, 2$, and the assumed sampling rate is $\frac{1}{\Delta t} = 120$ Hz. The driving inputs are independent white Gaussian noises with zero-mean and unit-variance. The frequency-domain causality is computed for each system using the single-model unconditional approach with the true parameter values, and the time-domain causality is computed by numerical integration of the frequency-domain causality over frequency. The left column of Figure 5-1 shows the true spectra of the transmitters and receivers for the different systems. The right column shows the true frequency-domain causality values for the different systems. Each plot represents a single setting of the receiver resonance frequency, with the different transmitter frequencies. For instance, the top right plot shows the values for all five transmitter settings with the receiver resonance set at 10 Hz. Similarly, the middle right and bottom right plots show the values with the receiver resonance held fixed at 30 Hz and 50 Hz, respectively.

5.1.1 Receiver Independence: Unconditional Frequency-Domain GG-Causality Does Not Depend on Receiver Dynamics

Perhaps the most noticeable aspect of Figure 5-1 is that the causality plots of the right column are identical, indicating that the frequency-domain causality values are independent of the resonance frequency of the receiver. This suggests that the unconditional causality is actually independent of the receiver dynamics. This can be seen explicitly by writing the causality function in terms of the VAR system components. The transfer function of the system is the inverse of the Fourier transform of the VAR equation,

$$H(\lambda) = \mathcal{A}^{-1}(\lambda) = \left(I - \sum_{p=1}^{P} A(p) \exp^{-ip\lambda} \right)^{-1}.$$

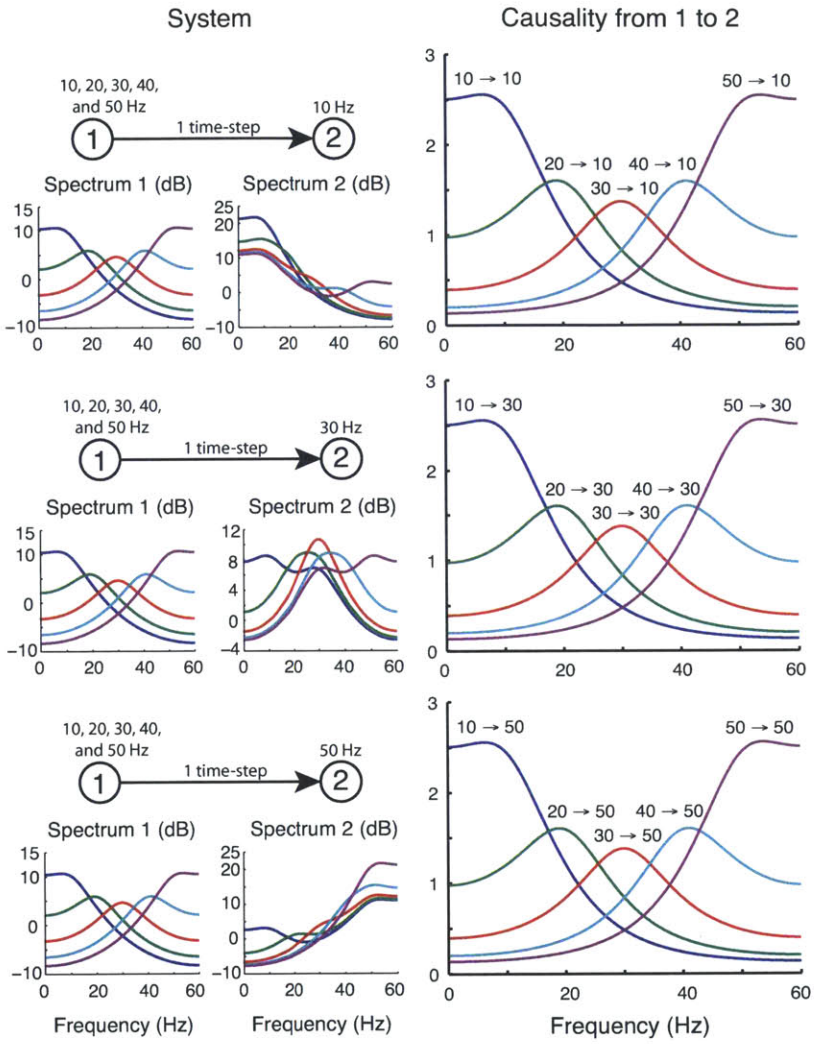

Figure 5-1: System and Unconditional Frequency-Domain Causality Values for Different Resonance Alignments of the VAR(2) Unidirectional Transmitter-Receiver Pair of Example 3. The channel is all-pass unit delay. Three receiver resonances are tested: 10 Hz, 30 Hz, and 50 Hz, shown in the top, middle, and bottom rows, respectively. For each setting of the receiver resonance, five transmitter resonances are tested: 10 Hz, 20 Hz, 30 Hz, 40 Hz, and 50 Hz. The left column shows the spectra. The right column shows the causality. As discussed in the text, the causality reflects the spectrum of the transmitter and channel, independent of the receiver dynamics. Hence, all three plots of the right column are identical and similar to the node 1 spectrum plots of the left column.

As shown in Appendix 5.3.1, the unconditional causality (Eqn. 2.2) from node 1 to node 2 can be expressed in terms of the AR matrices associated with each (possibly vector-valued) component

$$\begin{aligned} f_{x_1 \longrightarrow x_2}(\lambda) &= \ln \frac{\left| H_{2,2}(\lambda) \Sigma_{2,2} H_{2,2}^*(\lambda) + H_{2,1}(\lambda) \Sigma_{1,1} H_{2,1}^*(\lambda) \right|}{\left| H_{2,2}(\lambda) \Sigma_{2,2} H_{2,2}^*(\lambda) \right|} \\ &= \ln \frac{\left| \Sigma_{2,2} + \mathcal{A}_{2,1}(\lambda) \mathcal{A}_{1,1}^{-1}(\lambda) \Sigma_{1,1} \mathcal{A}_{1,1}^{-1*}(\lambda) \mathcal{A}_{2,1}^*(\lambda) \right|}{\left| \Sigma_{2,2} \right|}. \end{aligned} \tag{5.1}$$

Note how the $\mathcal{A}_{2,2}(\lambda)$ system component, and hence the $\{A_{2,2}(p)\}$ system parameters, which characterize the receiver dynamics, are absent. Thus, the unconditional frequency-domain causality does not reflect receiver dynamics.

The causality remains a function of the transmitter and channel dynamics and the input variances. In Figure 5-1, the causality is seen to parallel the transmitter spectrum, with the causality peak located at the transmitter resonance frequency. This makes sense, as the channel remains fixed as an all-pass unit delay, so the causality is determined by the transmitter spectrum, $\mathcal{A}_{1,1}^{-1}(\lambda) \Sigma_{1,1} \mathcal{A}_{1,1}^{-1*}(\lambda)$, in the second term of the numerator in Eqn. 5.1.

Even in this simple bivariate example the causality analysis is potentially misleading and could undermine comparisons between related systems. For example, when comparing the various transmitter settings when the receiver resonance is held at 30 Hz, the peak output power of the receiver is maximal when the transmitter resonance is aligned at 30 Hz (Figure 5-1, middle row, center plot). However, in this configuration, the causality is actually minimized (Figure 5-1, middle row, right plot). Because the causality is independent of the receiver dynamics, the size of the effect at the output is irrelevant. If the analyst's notion of what should be causal is in anyway related to the magnitude of response of an output variable, GG-causality will fail to characterize such system behavior. In such cases, application of GG-causality could be grossly misleading.

Similarly, the receiver independence property of GG-causality also impies that two systems can have identical causality functions, but completely different receiver dynamics. The frequency response and causality in this example is entirely reflective

of one AR parameter value, $A_{1,1}(1) = 2r_1 \cos(\theta_1)$. In a more general scenario, the dependence on the transmitter and channel VAR parameters can be highly complicated. Consequently, comparisons of causality values cannot be appropriately interpreted without first comparing the estimated VAR systems.

5.1.2 Receiver Independence: Unconditional and Conditional Time-domain and Conditional Frequency-Domain

The remaining forms of GG-causality—unconditional time-domain, conditional time-domain, and conditional frequency-domain—also appear to be independent of the receiver dynamics. For time-domain GG-causality, closed-form expressions were derived that illustrate receiver independence for both the unconditional and conditional cases. The details for these derivations are described in Appendix 5.3.2. Briefly, the time-domain GG-causality compares the prediction error variances from full and reduced models. The prediction error variance for the full model is simply the input noise variance, so the dependence of the GG-causality on the VAR parameters arises through the prediction error variance of the reduced model. In general, the reduced model can only be obtained numerically [7], which obscures the form of its dependence on the full model VAR parameters. Instead, a state-space representation of the VAR is used to derive an implicit expression for the reduced model prediction error variance in terms of the VAR parameters. It is then shown that the reduced model prediction error variance is independent of the receiver node VAR parameters. Thus, the time-domain GG-causality is independent of receiver dynamics. This is true for both the unconditional and conditional cases.

At present, for the reduced model in the conditional frequency-domain construct, I do not have a closed form expression with respect to the full model parameters. As a consequence, it cannot yet be determine analytically whether the conditional frequency-domain causality depends on the receiver node parameters. However, the relationship between the conditional time- and frequency-domain causality measures can be used to help analyze receiver dependence. Note that the frequency-domain

causality is a decomposition of the time-domain causality. In the unconditional case [26],

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f_{x_i \longrightarrow x_j}(\lambda)\, d\lambda \leq \mathcal{F}_{x_i \longrightarrow x_j},$$

with equality if the transmitter node is stable – i.e., iff the roots of

$$\det\left(I - \sum_{p=1}^{P} A_{i,i}(p)\, z^p\right)$$

are outside the unit circle – which is the case for the systems in Example 3. Similarly, in the conditional case [27],

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f_{x_i \longrightarrow x_j | x_k}(\lambda)\, d\lambda \leq \mathcal{F}_{x_i \longrightarrow x_j | x_k}, \tag{5.2}$$

with equality iff the roots of $\det\left(\sum_{p=1}^{\infty} G(p)\, z^p\right)$ are inside the unit circle, where G is the system function of Eqn. 2.4. Hence, the frequency-domain causality is a decomposition of the time-domain causality.

Since the time-domain conditional causality is independent of the receiver node, the relationship expressed in Eqn. 5.2 would, at a minimum, impose strict constraints on the form of any receiver dependence in the frequency-domain conditional causality. Example 1 from the Chapter 5 qualitatively agrees with this. As seen in Figure 3-2, the receiver dynamics do not appear to be reflected in the conditional frequency-domain causality. In the third column of Figure 3-2, the causality estimates from node 1 to node 2, the 10 Hz receiver resonance of node 2 appears to be absent. Similarly, in the causality estimates from node 2 to node 3, the 50 Hz receiver resonance of node 3 appears to be absent. Thus, the conditional frequency-domain causality appears to be independent of the receiver dynamics, similar to the other cases considered above.

Example 1 also helps to illustrate the challenge in interpreting GG-causality when both the transmitter and channel have dynamic properties (i.e., a frequency response). From node 1 to 2, the causality (top plot of column 3 of Figure 3-2) primarily reflects the spectrum of the transmitter, showing a prominent peak at 40 Hz (node 1, Figure 3-1). In contrast, the causality from node 2 to node 3 shows a nadir at approximately 15

Hz (middle plot from column 3 of Figure 3-2), which resembles the channel dynamics more than the node 2 transmitter dynamics (Figure 3-1). This example reinforces the earlier observation that causality values cannot be appropriately interpreted without examining the estimated system components.

5.2 Discussion

5.2.1 Summary of Results

Chapters 3 and 5 analyzed several problems encountered during Granger causality analysis, focusing primarily on the Geweke frequency-domain measure and the case where the generative processes were finite VARs of known order. I found that:

1. The practicalities of computing the conditional causality, particularly the use of separate VAR model fits and finite order reduced models, lead to model order selection issues. Because subsets of VAR processes are generally VARMA, there is a tradeoff between the bias of using the true full model order and the variance associated with a higher model order. Further, the use of separate model fits can lead to high sensitivity of the causality to uncertainties in model parameter estimates, producing spurious peaks and valleys, and even negative values, in the frequency domain.

2. The causality results, even for the simplest of examples, can be difficult to interpret. The causality appears to be independent of the receiver dynamics, and reflects a combination of the transmitter and channel dynamics. Additionally, the causality estimates are not informative of the system behavior without consideration of the full model estimate, and can be counterintuitive to one's notion of causality.

The problems stemming from practical computational considerations have been acknowledged to varying degrees, and have somewhat limited the adoption of the conditional measures. While it is generally known that VAR subsets are VARMA

and that conditional GG-causality is defined in terms of infinite time series histories, the impact of selecting a finite model order has not been studied carefully.

Model order selection can be problematic even in (single-model) parametric spectral estimation, where high model order can produce artifactual spectral peaks and shift estimated peak locations away from their true values. But in conditional GG-causality a full model spectrum must be "compared" to that of a reduced model. And while truncating the models at the true full model order poorly approximates the spectra under the reduced model, biasing the causality estimates, increasing the model order increases the uncertainty and the noisiness of the estimates.

The negative causality estimates and spurious peaks in the frequency domain stemming from separate model fits has been recognized previously [16]. This has led to proposed modifications of the conditional GG-causality computation that avoid these complications by relying on a single full model estimate [16, 7]. The analysis of Chapter 3 suggests that this additional sensitivity from the mismatch between the separate model fits is overshadowed by the more general bias and variance trade-off. While negative values can be eliminated by single model estimates, spurious peaks come about, not only through a mismatch between separate model fits, but also through two related sources – one due to bias in the reduced model estimate when a low model order is chosen, and another due to VAR parameter variance when a higher model order is chosen. The implications of this insight for the role of causality analysis in neuroscience, as part of a modeling and system identification framework, will be discussed below and in the following chapter.

Chapter 4 presented a novel method to correctly compute the conditional GG-causality. The method uses a state-space form of the full model VAR to obtain the correct reduced model via spectral factorization. This method avoids the computational problems of the standard method using separate model fits.

The problems of interpreting causality values are more fundamental. That GG-causality is unaffected by and not reflective of the receiver dynamics is entirely understandable from the original principle and definition of Granger causality, that of improved prediction. Receiver independence for a scalar, unconditional frequency-

domain example was previously noted in [31]. The analysis of this chapter shows that this holds more generally, for the vector, unconditional frequency domain, the unconditional time-domain, and conditional time-domain cases. The results also suggest that receiver independence may hold in the conditional frequency-domain case as well. This illustrates how use of GG-causality may be troublesome in neuroscience investigations – that the concept of causality defined by Granger may differ from the notion of causality of interest to the analyst, particularly if that notion encompasses some aspect of the effect on an output variable.

5.2.2 Implications for Other Causality Approaches

The utility of the GG-causality, and causality analyses in general, hinge on the notions of causality involved and the model class underlying such methods. There are three notions of causality to consider: the formal notion of causality defining the causality analysis approach, the conceptual notion of causality of interest to the analyst, and finally the practical causality values computed in the causality estimation procedure. It is imperative that these three notions be in agreement. A variety of alternative causality measures beyond GG-causality have been proposed to analyze neuroscience data. Each of these methods reflect different formal and conceptual notions of causality, with estimation procedures whose properties vary.

Two modifications to the method for computing conditional GG-causality have been proposed, seeking to eliminate the need for separate model fits and the subsequent sensitivity issues – spurious peaks and valleys and negative values. Chen et al. [16] suggested using a reduced model with colored noise formed from the components of the estimated full model. The rotation used in [27] to eliminate the instantaneous causality (See Appendix 2.4.2) is applied at each frequency. This rotated transfer function is then used in Eqn. 2.3. However, this approach does not use the correct decomposition for the reduced process. Barnett and Seth [7] propose fitting the reduced model and using it to directly compute the spectral components of Eqn. 2.5. The method appears to avoid the additional sensitivity of the separate model fits, but in order to eliminate bias in the reduced model estimate, uses a large order, and

is thus susceptible to the resulting increase in variance. Both of these methods also still face the interpretation problem.

Several alternative measures of frequency-domain causality, all based on VAR models, have also been proposed. These include the directed transfer function (DTF) [34, 35], the coherence, the partial coherence, the direct directed transfer function (dDTF) [39, 38], and the partial directed coherence (PDC) [5]. Each of these measures, with the exception of the dDTF, is based on a single model estimate/computation, and so do not suffer from the sensitivity or model subset issues of the Granger-Geweke approach. However, these measures do face an analogous interpretation problem to that demonstrated for the Geweke measure, specifically in understanding how the causality values relate to the behaviors of the system components.

Each of the causality measures represents a different functional combination of the properties of the individual system components, taking on different values in the case of true causal connection. Consequently, they each represent conceptually distinct notions of causality, capturing different aspects of the system behavior. For instance, consider a truly causal connection from a transmitter to a receiver within a larger system with conditional nodes. The DTF represents the transfer function, not from the transmitter to the receiver, but instead from an exogenous white noise input at the transmitter to the receiver. Depending on the normalization used, it may actually include the dynamics of all systems components. It essentially captures all information passing through the transmitter *directed* to the receiver, but this information flow is *not direct*. The PDC, on the other hand, represents the actual channel between the nodes, but (depending on normalization) neglects both the spectral content being transmitted and the response of the receiver. And both of these measures differ from the Granger-Geweke notion of improved prediction, which captures information reaching the receiver, i.e. the spectrum of the transmitter as passed through the channel, but not the receiver response. Unfortunately, these causality approaches are often discussed as interchangeable alternatives [21, 3], but clearly they reflect different notions of causality, and different aspects of the underlying system.

In practice, many causality analyses focus primarily on hypothesis testing for a

direct connection between nodes, without regard to the values or functional form of the underlying causality measures computed along the way. However, the direct connection between nodes is simply the corresponding cross-components of the VAR parameter matrices, and absence of connection corresponds to this set of parameters being all equal to zero. This condition is directly and easily tested for by the 1-step detection test, as described in [42]. However, this detection test has not been, to my knowledge, used in neuroscience applications, despite its potential advantages in computational and statistical efficiency compared to more widely-applied permutation-based tests on causality measures. The causality measures discussed above agree on this instance of non-causality, taking on zero values when the channel parameters are zero.

While bivariate GG-causality agrees on the above non-causality condition, the general condition for Granger non-causality in VAR models is more complicated, corresponding to an ∞-step detection test [18], as described in [42]. In the case of a VAR with M nodes of order P, the (MP+1)-step test is sufficient to determine Granger non-causality and corresponds to a nonlinear set of restrictions on the set of VAR parameters [42]. This clarifies that GG-causality contains some indirect aspects of the system and is not entirely direct. As mentioned in Section 2.3, this is suggested by full model noise processes for the conditional nodes appearing in the conditional frequency-domain definition and the time-domain equivalence from which it derived. This is also indicated in Appendix 5.3.2, where it is seen that the conditional node dynamics are also eliminated from the time-domain GG-causality, but indirect channels and conditional node variances remain. The nonlinearity of the multi-step test makes the computation more difficult than the 1-step test and is also suggestive of the non-standard distribution of conditional GG-causality.

Like Granger-Geweke, each of these approaches is subject to the properties and limitations imposed by the underlying VAR model structure. That subsets of VAR processes are VARMA introduces a challenging model order selection problem for practical computation of GG-causality. In general, model order selection can be difficult even when a single VAR model is being estimated. AIC is a principled and

popular criterion for order selection, but it is an inconsistent estimator. Consistent estimators exist, but they tend to underestimate the order for small to moderate sized time series ($\sim$ 100 time points) [42]. Such slight discrepancies in order may not be consequential in the context of prediction, but in spectral estimation they can be of great import. Changes in model order alter the number and location of spectral peaks. If the wrong order is chosen, the resulting spectra, and subsequent causality analysis, could be highly misleading, particularly in data featuring prominent oscillations.

VAR models represent linear relationships within stationary, time-invariant time series data. This structure limits the generative system behavior that can be modeled. For example, linear stability severely restricts the possible feedback connections and frequency alignments allowed between nodes in systems with high dynamic-range oscillations. When applied to more general generative processes, it cannot represent nonlinear phenomena like cross-frequency coupling, observed frequently in neural oscillation data [32, 15]. Any notion of causality pertaining to such phenomena cannot be defined from VAR. More generally, application of VAR-based causality approaches to generative processes outside the VAR model class can lead to erroneous results [37, 25].

Causality approaches for more general systems have also been proposed. Transfer entropy methods [48, 6, 46] can be viewed as a generalization of the Granger approach, and are applicable to nonlinear systems. However, such methods require significant amounts of data, are computationally intensive, and present their own model identification issues, such as selection of the embedding order. These approaches are strictly time-domain measures, and their relationship to GG-causality [6] suggests that they face the same interpretation problems, including independence of receiver dynamics. Other causality approaches based on biophysical models, such as dynamic causal modeling (DCM) [23], have also been proposed. The models, computational methods, and notions of causality involved in such approaches are entirely different from the the time series-based causality approaches discussed here. The interpretation and statistical properties of these methods are also the subject of ongoing analysis [17, 40, 22, 41].

5.2.3 Implications for the Role of Causality Analysis in Neuroscience

Many of the problems and limitations of causality analysis discussed in this work arise more critically in the neuroscience context than in other applications. In econometrics, for example, the objective is to obtain parsimonious models for making predictions. In neuroscience, the aim is to understand the underlying physiology, in particular the frequency-specific oscillatory behaviors of neural circuit components and their interactions within the system. From this physiological systems perspective, the system model estimate is the fundamental element in the analysis, characterizing the behaviors of the individual components and their connections. Causality measures are therefore a secondary reformulation of information already contained in the model estimate. As discussed, they are only fully interpretable in the context of the system model estimate. The utility of causality methods for characterizing interactions between system components is ultimately limited by the ability of the underlying model class to adequately represent the dynamics of the generative physiological system. Thus, our understanding of causal relationships might best be advanced by developing better models and estimation methods that can more faithfully represent neurophysiological dynamics. This will be discussed further in Chapter 6.

5.3 Appendix

5.3.1 Receiver Independence of Unconditional Frequency- Domain GG-Causality

This appendix shows that the unconditional frequency-domain GG-causality is independent of the receiver dynamics. Without loss of generality, receiver independence is demonstrated for the causality from transmitter x_1 to receiver x_2, i.e. the causality is independent of the parameters governing the internal receiver dynamics $\{A_{22}(p)\}$. First, the full model transfer function $H(\lambda)$ is written in terms of block components. The transfer function is the inverse of the Fourier transform of the VAR,

$\mathcal{A}(\lambda) = I - \sum_{p=1}^{P} A(p) e^{-ip\lambda}$, so from the matrix inversion lemma

$$\begin{bmatrix} H_{11}(\lambda) & H_{12}(\lambda) \\ H_{21}(\lambda) & H_{22}(\lambda) \end{bmatrix} = \begin{bmatrix} \mathcal{A}_{11}(\lambda) & \mathcal{A}_{12}(\lambda) \\ \mathcal{A}_{21}(\lambda) & \mathcal{A}_{22}(\lambda) \end{bmatrix}^{-1}$$
$$= \begin{bmatrix} \mathcal{A}_{11}^{-1} + \mathcal{A}_{11}^{-1}\mathcal{A}_{12}\left(\mathcal{A}_{22} - \mathcal{A}_{21}\mathcal{A}_{11}^{-1}\mathcal{A}_{12}\right)^{-1}\mathcal{A}_{21}\mathcal{A}_{11}^{-1} & -\mathcal{A}_{11}^{-1}\mathcal{A}_{12}\left(\mathcal{A}_{22} - \mathcal{A}_{21}\mathcal{A}_{11}^{-1}\mathcal{A}_{12}\right)^{-1} \\ -\left(\mathcal{A}_{22} - \mathcal{A}_{21}\mathcal{A}_{11}^{-1}\mathcal{A}_{12}\right)^{-1}\mathcal{A}_{21}\mathcal{A}_{11}^{-1} & \left(\mathcal{A}_{22} - \mathcal{A}_{21}\mathcal{A}_{11}^{-1}\mathcal{A}_{12}\right)^{-1} \end{bmatrix}.$$

Writing the causality in terms of the the transfer function and noise covariance components,

$$f_{x_1 \longrightarrow x_2}(\lambda) = \ln \frac{\left|H_{22}(\lambda)\Sigma_{22}H_{22}^{*}(\lambda) + H_{21}(\lambda)\Sigma_{11}H_{21}^{*}(\lambda)\right|}{\left|H_{22}(\lambda)\Sigma_{22}H_{22}^{*}(\lambda)\right|}.$$

Substituting in terms of $\mathcal{A}$ components from above,

$$f_{x_1 \longrightarrow x_2}(\lambda)$$
$$= \ln \frac{\left|\left(\mathcal{A}_{22} - \mathcal{A}_{21}\mathcal{A}_{11}^{-1}\mathcal{A}_{12}\right)^{-1}\left\{\Sigma_{2,2} + \mathcal{A}_{21}\mathcal{A}_{11}^{-1}\Sigma_{11}\left(\mathcal{A}_{11}^{-1}\right)^{*}\mathcal{A}_{21}^{*}\right\}\left(\left(\mathcal{A}_{22} - \mathcal{A}_{21}\mathcal{A}_{11}^{-1}\mathcal{A}_{12}\right)^{-1}\right)^{*}\right|}{\left|\left(\mathcal{A}_{22} - \mathcal{A}_{21}\mathcal{A}_{11}^{-1}\mathcal{A}_{12}\right)^{-1}\Sigma_{22}\left(\left(\mathcal{A}_{22} - \mathcal{A}_{21}\mathcal{A}_{11}^{-1}\mathcal{A}_{12}\right)^{-1}\right)^{*}\right|}.$$

Recall that for a matrix $|\Sigma| = \sqrt{\det \Sigma'\Sigma}$. Because for square matrices the determinant of the product is equal to the product of the determinants, the determinants of all factors in both the numerator and denominator separate. Consequently, the determinants of the first factors, $\left(\mathcal{A}_{22} - \mathcal{A}_{21}\mathcal{A}_{11}^{-1}\mathcal{A}_{12}\right)^{-1}$, the last factors, $\left(\left(\mathcal{A}_{22} - \mathcal{A}_{21}\mathcal{A}_{11}^{-1}\mathcal{A}_{12}\right)^{-1}\right)^{*}$, and their corresponding transposes, cancel from the numerator and denominator. This results in the final equality,

$$f_{x_1 \longrightarrow x_2}(\lambda) = \ln \frac{\left|\Sigma_{22} + \mathcal{A}_{21}(\lambda)\mathcal{A}_{11}^{-1}(\lambda)\Sigma_{11}\mathcal{A}_{11}^{-1*}(\lambda)\mathcal{A}_{21}^{*}(\lambda)\right|}{\left|\Sigma_{22}\right|}.$$

Thus, the receiver dynamics $\mathcal{A}_{22}(\lambda)$ do not appear. Because the time-domain causality is the integral of the frequency-domain causality, the time-domain unconditional GG-causality is independent of the receiver dynamics as well.

5.3.2 Receiver Independence of Time-Domain GG-Causality

This appendix shows that the general time-domain GG-causality is independent of the receiver dynamics. In particular, receiver independence holds for the conditional time-domain GG-causality. This is demonstrated using the state-space form, from which the computation using spectral factorization was derived in Chapter 4.

The GG-causality from node i to node j compares the 1-step prediction-error variance of the full model, Σ^f_{jj}, with the 1-step prediction-error variance of the reduced model, $\Sigma^{r(i)}_{jj}$,

$$\mathcal{F}_{x_i \longrightarrow x_j} = \ln \frac{|\text{var}\left(x_{j,t} - \hat{x}^{r(i)}_{j,t}\right)|}{|\text{var}\left(x_{j,t} - \hat{x}^{f}_{j,t}\right)|} = \ln \frac{|\Sigma^{r(i)}_{jj}|}{|\Sigma^{f}_{jj}|}.$$

The prediction-error covariance of the full model is just the input noise covariance $\Sigma^f = \Sigma_w$, so the dependence of the GG-causality on the VAR parameters arises through the prediction-error covariance of the reduced model. A closed form expression for the reduced model with respect to the full model parameters is yet unknown. In general, the reduced model can only be obtained numerically [7], which obscures the form of its dependence on the full model VAR parameters. Instead, the state-space representation of the VAR can be used to derive an expression for the reduced model prediction-error variance in terms of the VAR parameters. From this, it can be seen that the receiver node VAR parameters do not enter into the equation for the reduced model prediction-error variance, and thus, the time-domain GG-causality is independent of receiver dynamics. This is true for both the unconditional and conditional cases.

To put the model in state-space form, write the VAR(P) full model as the augmented

VAR(1) state equation,

$$\begin{aligned}\overline{x}_t =& F\overline{x}_{t-1} + Gw_t \\ =& \begin{bmatrix} A^f(1) & \cdots & A^f(P-1) & A^f(P) \\ & \mathrm{I} & & 0 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ \vdots \\ x_{t-p} \end{bmatrix} + \begin{bmatrix} \mathrm{I} \\ 0 \end{bmatrix} w_t^f,\end{aligned} \tag{5.3}$$

where $\overline{x}_t = \begin{bmatrix} x_t^T & x_{t-1}^T & \dots & x_{t-(P-1)}^T \end{bmatrix}^T$ is the augmented state. The observation equation is

$$\begin{aligned} y_t =& H\overline{x}_t + v_t \\ =& \begin{bmatrix} \begin{pmatrix} \mathrm{I} & 0 \end{pmatrix} & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} x_t \\ x_{t-1} \\ \vdots \\ x_{t-(P-1)} \end{bmatrix} = x_{j,t} \end{aligned} \tag{5.4}$$

where H is a selection matrix that selects the observed components from the top block of the state, the observation noise is zero, $v_t = 0$, and by convention the observed/effect components are ordered first in each block, $x_t = \begin{bmatrix} x_{j,t}^T & x_{i,t}^T \end{bmatrix}^T$.

The reduced model prediction-error covariance is simply the observed components of the state prediction-error covariance,

$$\Sigma^{r(i)} = H\Sigma_{t+1|t}H^T. \tag{5.5}$$

The state prediction-error covariance and state (filtered)-error covariance are given recursively by the standard Kalman filter equations,

$$\Sigma_{t+1|t} = F\Sigma_{t|t}F^T + G\Sigma_w G^T \tag{5.6}$$

and

$$\Sigma_{t|t} = \Sigma_{t|t-1} - \Sigma_{t|t-1} H^T \left(H \Sigma_{t|t-1} H^T + R \right)^{-1} H \Sigma_{t|t-1}, \tag{5.7}$$

where in this case the observation noise covariance is zero, $R = 0$. Designate the blocks of the prediction-error and state covariances by

$$\Sigma_{t+1|t} = \begin{bmatrix} K_{t+1} & L_{t+1} \\ L_{t+1}^T & N_{t+1} \end{bmatrix} \quad \text{and} \quad \Sigma_{t|t} = \begin{bmatrix} B_t & C_t \\ C_t^T & D_t \end{bmatrix}, \quad \text{respectively.}$$

To see the VAR parameter dependence, write out the state prediction-error equation, Eqn. 5.6, by components. To demonstrate without loss of generality, a VAR(2) is used,

$$\Sigma_{t+1|t} = \begin{bmatrix} A^{(1)} & A^{(2)} \\ \mathrm{I} & 0 \end{bmatrix} \begin{bmatrix} B_t & C_t \\ C_t^T & D_t \end{bmatrix} \begin{bmatrix} A^{(1)T} & \mathrm{I} \\ A^{(2)T} & 0 \end{bmatrix} + \begin{bmatrix} \mathrm{I} \\ 0 \end{bmatrix} \Sigma_w \begin{bmatrix} \mathrm{I} & 0 \end{bmatrix}. \tag{5.8}$$

In this case, the blocks K, L, N, B, C, and D all have dimensions $M \times M$, the same dimensions as the VAR matrices $A^{(p)}$, where M is the number of system components. For notational simplicity, the superscript f from the VAR matrices are omitted and the AR lag index is moved to superscript. Because the reduced components are directly observed, $y_t = x_{j,t}$, the associated state-error variances are zero, so the only non-zero components of the blocks are the covariances of the unobserved/omitted components,

$$B_t = \begin{pmatrix} 0 & 0 \\ 0 & \sigma^2_{(i)t,t|t} \end{pmatrix}, \quad C_t = \begin{pmatrix} 0 & 0 \\ 0 & \sigma^2_{(i)t,t-1|t} \end{pmatrix}, \quad \text{and} \quad D_t = \begin{pmatrix} 0 & 0 \\ 0 & \sigma^2_{(i)t-1,t-1|t} \end{pmatrix}.$$

Carrying out the matrix multiplication of Eqn. 5.8 for the prediction-error covariance,

$$\Sigma_{t+1|t} = \begin{bmatrix} K_{t+1} & L_{t+1} \\ L_{t+1}^T & N_{t+1} \end{bmatrix}$$
$$= \begin{bmatrix} \Sigma_w + A^{(1)} B_t A^{(1)T} + A^{(2)} C_t^T A^{(1)T} + A^{(1)} C_t A^{(2)T} + A^{(2)} D_t A^{(2)T} & A^{(1)} B_t + A^{(2)} C_t^T \\ B_t A^{(1)T} + C_t A^{(2)T} & B_t \end{bmatrix}.$$

The top left block of $\Sigma_{t+1|t}$ is

$$K_{t+1} = \Sigma_w + A^{(1)} B_t A^{(1)T} + A^{(2)} C_t^T A^{(1)T} + A^{(1)} C_t A^{(2)T} + A^{(2)} D_t A^{(2)T}$$

$$= \begin{pmatrix} \Sigma_{w_j} + A_{ji}^{(1)} \sigma_{t,t|t}^2 A_{ji}^{(1)T} + A_{ji}^{(2)} \sigma_{t-1,t|t}^2 A_{ji}^{(1)T} & A_{ji}^{(1)} \sigma_{t,t|t}^2 A_{ii}^{(1)T} + A_{ji}^{(2)} \sigma_{t-1,t|t}^2 A_{ii}^{(1)T} \\ +A_{ji}^{(1)} \sigma_{t,t-1|t}^2 A_{ji}^{(2)T} + A_{ji}^{(2)} \sigma_{t-1,t-1|t}^2 A_{ji}^{(2)T} & +A_{ji}^{(1)} \sigma_{t,t-1|t}^2 A_{ii}^{(2)T} + A_{ji}^{(2)} \sigma_{t-1,t-1|t}^2 A_{ii}^{(2)T} \\ A_{ii}^{(1)} \sigma_{t,t|t}^2 A_{ji}^{(1)T} + A_{ii}^{(2)} \sigma_{t-1,t|t}^2 A_{ji}^{(1)T} & \Sigma_{w_i} + A_{ii}^{(1)} \sigma_{t,t|t}^2 A_{ii}^{(1)T} + A_{ii}^{(2)} \sigma_{t-1,t|t}^2 A_{ii}^{(1)T} \\ +A_{ii}^{(1)} \sigma_{t,t-1|t}^2 A_{ji}^{(2)T} + A_{ii}^{(2)} \sigma_{t-1,t-1|t}^2 A_{ji}^{(2)T} & +A_{ii}^{(1)} \sigma_{t,t-1|t}^2 A_{ii}^{(2)T} + A_{ii}^{(2)} \sigma_{t-1,t-1|t}^2 A_{ii}^{(2)T} \end{pmatrix}. \tag{5.9}$$

The upper right (and transpose of the lower left) block of $\Sigma_{t+1|t}$ is

$$L_{t+1} = A^{(1)} B_t + A^{(2)} C_t^T = \begin{pmatrix} 0 & A_{ji}^{(1)} \sigma_{t,t|t}^2 + A_{ji}^{(2)} \sigma_{t-1,t|t}^2 \\ 0 & A_{ii}^{(1)} \sigma_{t,t|t}^2 + A_{ii}^{(2)} \sigma_{t-1,t|t}^2 \end{pmatrix}. \tag{5.10}$$

The lower right block of $\Sigma_{t+1|t}$ is

$$N_{t+1} = B_t = \begin{pmatrix} 0 & 0 \\ 0 & \sigma_{t,t|t}^2 \end{pmatrix}. \tag{5.11}$$

Stepping forward the Kalman filter with Eqn. 5.7, the next state-error covariance is

$$\Sigma_{t+1|t+1} = \Sigma_{t+1|t} - \Sigma_{t+1|t} H^T \left(H \Sigma_{t+1|t} H^T \right)^{-1} H \Sigma_{t+1|t}. \tag{5.12}$$

Expanding,

$$
\begin{bmatrix} B_{t+1} & C_{t+1} \\ C_{t+1}^T & D_{t+1} \end{bmatrix}
$$

$$
= \begin{bmatrix} K_{t+1} & L_{t+1} \\ L_{t+1}^T & N_{t+1} \end{bmatrix} - \begin{bmatrix} K_{t+1} & L_{t+1} \\ L_{t+1}^T & N_{t+1} \end{bmatrix} \left[\begin{array}{c} \mathrm{I} \\ 0 \\ \hline 0 \\ 0 \end{array}\right] (K_{t+1,jj})^{-1} \left[\begin{array}{cc|cc} \mathrm{I} & 0 & 0 & 0 \end{array}\right] \begin{bmatrix} K_{t+1} & L_{t+1} \\ L_{t+1}^T & N_{t+1} \end{bmatrix}
$$

$$
= \begin{bmatrix} \begin{pmatrix} 0 & 0 \\ 0 & K_{t+1,ii} - K_{t+1,ij}K_{t+1,jj}^{-1}K_{t+1,ji} \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & L_{t+1,ii} - K_{t+1,ij}K_{t+1,jj}^{-1}L_{t+1,ji} \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \\ 0 & L_{t+1,ii}^T - L_{t+1,ji}^T K_{t+1,jj}^{-1}K_{t+1,ji} \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & N_{t+1,ii} - L_{t+1,ji}^T K_{t+1,jj}^{-1}L_{t+1,ji} \end{pmatrix} \end{bmatrix}.
$$

It is seen that the state-covariance structure is preserved, as expected, and that the covariances of the unobserved components are given iteratively by

$$
\begin{aligned}
\sigma^2_{(i)t+1,t+1|t+1} &= K_{t+1,ii} - K_{t+1,ij}K_{t+1,jj}^{-1}K_{t+1,ji} \\
\sigma^2_{(i)t+1,t|t+1} &= L_{t+1,ii} - K_{t+1,ij}K_{t+1,jj}^{-1}L_{t+1,ji} \\
\sigma^2_{(i)t,t|t+1} &= N_{t+1,ii} - L_{t+1,ji}^T K_{t+1,jj}^{-1}L_{t+1,ji}.
\end{aligned} \tag{5.13}
$$

The state prediction- and filtered-error covariances of Eqns. 5.6 and 5.7 have been consolidated to Eqns. 5.9, 5.10, 5.11, and 5.13. From this set of equations, it is seen that the parameters determining the receiver dynamics, $\{A_{jj}^{(p)}\}$, do not appear. These equations can be computed recursively, and in steady-state, they form a set of non-linear equations, the solution of which determines the reduced model prediction-error variance. Hence, because the $\{A_{jj}^{(p)}\}$ parameters do not appear, the reduced model prediction-error variance and the time-domain causality is, as seen before in the unconditional frequency-domain, independent of the receiver dynamics.

In fact, it is also apparent that receiver independence holds for the conditional case, as well, and that the causality is further independent of the dynamics of the conditional nodes. This is because any conditional time series x_k would be included with x_j as directly observed components of the reduced model. Thus, the causality is only a function (a very complicated function) of the channel parameters $\{A_{ji}^{(p)}\}$, the transmitter parameters $\{A_{ii}^{(p)}\}$, and the input noise covariances Σ_w.

Chapter 6

Causality and Modeling

This chapter expands on the discussion of the previous chapter. It explores how causality questions are approached in other fields—econometric time series, classical regression, and probabilistic reasoning. Though the methods and perspectives of these fields can differ greatly from those of neuroscience, the comparison is informative of the fundamental difficulties underlying investigations of neuroscientific causality. The various neuroscience causality measures mentioned in the last chapter are analogous to various causality statements in these other fields. However, as will be discussed, the careful considerations of the data, the model, and the interpretations of parameters that these methods require are largely absent in their application in neuroscience. With this wider view, the neuroscience perspective of causality is revisited and a framework for addressing questions of causality is placed in the broader context of modeling and system identification.

6.1 Time Series Perspective

The view of causality in time series analysis, particularly the field of econometrics, from which the Granger concept of causality and the Geweke causality measures emerged, differs substantially from the direction neuroscience has gone. Firstly, causality measures, even those of Geweke, are rarely, if ever, used. Instead, restriction tests on sets of the VAR parameters are employed. To detect a direct connection from

$\{x_{i,t}\}$ to $\{x_{j,t}\}$ conditional on $\{x_{k,t}\}$ in a VAR(P) model, one simply tests against the null hypothesis

$$H_0: \quad A_{j,i}(p) = 0 \quad \forall p \in \{1, \ldots, P\}.$$

This test of direct connection amounts to a test of improvement in 1-step prediction, and is simply a likelihood ratio test between the VAR model with and without the connection.

One can also test for improvement in h-step prediction. If, as in the previous state-space formulations of Chapters 4 and 5, F is the VAR(1) form of the VAR(P) model and U^f is a selection matrix that selects the top block-row, such that for the h-th transition of the model

$$U^f F^h = \begin{bmatrix} A^h(1) & \cdots & A^h(P) \end{bmatrix},$$

then the test for h-step connection is against the null hypothesis

$$H_0: \quad A^s_{j,i}(p) = 0 \quad \forall p \in \{1, \ldots, P\}, \quad \forall s \leq h.$$

This h-step test determines whether information from $\{x_{i,t}\}$ can reach $\{x_{j,t}\}$ in at least h time steps. It tests whether the cross-parameters from i to j for all lags up to h transitions are zero. This is a nonlinear set of restrictions on the VAR parameters $\{A(p)\}$ and is no longer a likelihood ratio test between nested models. It is also not a test of direct connection.

In general, the Granger concept of causality corresponds to an ∞-step test above. Dufour and Renault [18] showed that in the case of a VAR(P) model of M time series, the test against Granger non-causality is given by a $(MP+1)$-step test for improvement of prediction. This is further indication that conditional GG-causality and the Granger concept of improvement of prediction does not assess direct connection within a system. The nonlinearity of the parameter restrictions also point to the non-standard distribution of conditional GG-causality estimates.

Second, to the degree that Granger causality is assessed in econometrics, it is

generally used more for exploratory analysis—to evaluate possible input variables to an output variable of interest and/or to decide whether to include an input variable in an analysis. This use of the Granger concept of "improvement of prediction" as a criterion for building models may seem odd in that it breaks from the traditional regression modeling approach of building up a model by comparing nested models to test for significance of new sets of parameters. However, in the econometric context, the possible value in the Granger concept is understandable. Very often, the modeling objective—that is, the intended use of the model—is prediction. From this perspective, both accuracy and precision are paramount. And while a detailed model of the system with availability of data for all variables would be ideal, from a practical standpoint the correspondence between the model and system is irrelevant.

Lastly, while the generative system is of limited interest, the responses of variables to perturbations is still important, for example in economic policy. Consequently, though causality measures are rarely used, especially not the frequency-domain measures, other structural analysis methods have been developed, such as impulse response analyses and forecast error variance decompositions.

Impulse response analysis examines the MA form of the VAR model. In econometrics, the view is often that the variables reflect a system in equilibrium about its mean and the question of interest is in the change of one variable due to an external shock in another. The i-th column of the p-lag MA matrix, $\Phi(p) = U^f F^p \left(U^f\right)^T$, represents the reaction of the variables to a unit shock in x_i, p time points ago. The accumulated response to a shock is also of potential interest and is given by $\Psi(n) = \sum_{p=0}^{n} \Phi(p)$, the n-th interim multiplier.

Depending on the desired interpretation of the analyst, modifications to impulse response analysis may be employed. For example, the unit impulses above do not reflect the scale of the perturbations, $\{w_t\}$, so an alternative form scales the shock by the standard deviations. More generally, if the noise covariance matrix Σ_w is not diagonal, then shocks in the variables are instantaneously correlated, so instead the responses to orthogonal impulses can be analyzed by rotating the impulse response matrices, $\{\Theta(p)\} = \{\Phi(p) D\}$, where D is the lower triangular factorization of the

noise covariance $\Sigma_w = DD^T$. This is the same rotation as that used to remove the instantaneous causality prior to the computation of frequency-domain GG-causality, and as in that case, the responses to orthogonal impulses are dependent on the order of the components of the vector time series. This order is now crucial and must be specified based on a priori knowledge. With orthogonalized impulses, yet another analysis tool is forecast error variance decomposition, which computes the proportion of the h-step prediction-error variance of one time series due to an orthogonal impulse in another. Other decompositions of the noise covariance are possible as well, but they correspond to other types of shocks, which require their own interpretation.

Impulse response analysis is the time-domain version of the DTF. While in neuroscience the DTF is sometimes misinterpreted as causality between the variables, in econometrics it is understood to reflect response to external shocks. But even with this clarification, there remains ambiguity in such analyses, requiring further specification, namely how to interpret the shocks. Thus, while the adoption of time series methods in econometrics allowed flexible modeling of data without a priori knowledge of a system of equations, the analysis and interpretation of a model requires its own restrictive assumptions.

The many structural analysis methods in econometrics are certainly employable for analyzing time series models of neuroscience data, but the accompanying interpretations and the underlying scientific views and objectives differ greatly from the interests of neuroscience investigators. The most important point from the econometric view that is applicable to causality analysis in neuroscience is that causality analysis and other structural analyses are specific statements of chosen properties of a chosen model. The validity of the results then depends largely on the appropriateness of the chosen model and method in terms of representing the system properties and answering the scientific question of interest.

Most of the points of this section come from [42]. Refer there and the references therein for further details of econometric time series, more general time series model classes, and additional complications that arise in causality analyses.

6.2 Classical Regression Perspective

As in time series analysis, classical regression typically addresses prediction of an output variable y based on input variables (predictors) x. Regression can also be used to make statements of causality, such as what would happen to y as a result of treatment variable T? The primary difficulty is that only the outcome of one treatment can be observed, so instead of directly computing the treatment effect, one must use other means, such as experimental randomization and/or statistical modeling adjustments, to compare the potential outcomes between similar units with randomly assigned treatments. In the case of binary treatments T^0 and T^1, with potential outcomes y_i^0 and y_i^1, the treatment effect, $y_i^1 - y_i^0$, is given by the regression coefficient of T.

However, as in time series analysis, there are fundamental differences between regression and causal interpretation, and interpretations of causality require much stricter assumptions. Improvement in prediction does not necessarily equal causality. Differences in treatment and control groups can result in treatments with no causal effect producing improved prediction or, conversely, treatments with a significant causal effect producing similar predictive distributions for treatment and control. To overcome this, all relevant pre-treatment variables must be controlled for by including them in the regression. The regression coefficient of the treatment variable can then be considered the causal effect.

Additional assumptions and characteristics of the data must also be considered. One crucial assumption is the ignorability assumption—that the probability of treatment assignment is equal conditional on the covariates. Otherwise, the treatment assignment would have to be modeled. In observational studies, where ignorability can't be proven, this is addressed by including as much covariate information as possible. Then, one must also consider the range and distribution of the data. Estimating causal effects requires overlap and balance of the control and treatment distributions conditional on the confounding variables. Results may not generalize to values of the variables insufficiently supported by the data, where the model is extrapolating or

biased due to imbalance.

Even if the data satisfy the above assumptions, causality interpretations still rest on the validity of the model, which is determined in large part by a priori understanding of the mechanism of the treatment effect and the relationship between the relevant variables.

Thus, one must yet consider the specificity of the causal statement being made. For example, whether the treatment is directly applied or merely made available as a replacement or a supplement for the control. Whether the treatment considered actually contains separate treatment aspects, which may be difficult to disentangle. Or whether there are other possible control conditions. Some desired causality statements are not well defined, such as "What is the strongest cause of an output from a multitude of predictor variables?" because the actual causal effect being sought is not specified.

The methods and concepts of causal inference in classical regression are primarily aimed at causality statements of single-valued variables, not the signals and systems of functional neuroscience data, and are thus not directly applicable. In fact, whereas determination and understanding of the system is the objective of causality analysis in neuroscience, the system, particularly the possible treatment mechanism, is assumed known, or at least acknowledged to be determined by choice of model structure. The principles applicable to neuroscience are the care and consideration of the necessary assumptions, the dependence on the model, and the specificity of the causal statement being made. Different model choices, assumptions, and interpretations lead to different causality statements.

For further details of causal interpretation in regression see [24] and references therein.

6.3 Probabilistic Reasoning Perspective

This section presents a few elements from the probabilistic reasoning view of causality, specifically from the framework advocated by Pearl [45]. According to the Pearl

framework, the system is defined entirely a priori. A structural model of the underlying system is formed by specifying the functional dependencies between the variables from existing knowledge and understanding of their relationships. From this model, various quantitative statements of causality can be made. These statements of causality reflect different sets of actions of the components within the system, different circumstances surrounding the operation and observation of the system, and inevitably the philosophical view of the investigator.

The range of practical applications is extensive, from artificial intelligence and natural language processing to medical and legal decision making. Pearl [45] is primarily a philosophical work, aimed at constructing the machinery—the definitions, structures, and methods—of a causality framework that corrects inconsistencies known to arise in the more classical approaches of logic and philosophy. The methods developed and advocated by Pearl are mostly formulated and demonstrated in terms of boolean variables. Extension to discrete and real valued variables and incorporation of probabilistic uncertainties is relatively straightforward. However, this framework is not easily adapted to signals and systems as functions of time. As such, the methods of Pearl are not directly applicable to many neuroscience applications.

In fact, the often intended use of causality analysis in neuroscience, the identification of the system, is not considered at all in the probabilistic reasoning view—it is taken as a given. Further, the discussion of various causality measures in neuroscience is seen to be a conflation of the philosophical quest for a universal framework for causality and the scientific interest in making specific statements about the properties of a system and finding robust methods to estimate those properties. Thus, the most important aspects of the probabilistic reasoning perspective applicable to neuroscience are conceptual—the many intuitive notions and concepts that underly various definitions of causality and that such definitions reflect different philosophical choices.

6.4 Back to Causality in Neuroscience

Perhaps the most important lesson gathered from the various perspectives above is that analysis of causality involves several separate components, and errors can occur at different levels of the analysis. One component establishes the system model and functional relationship between the variables. A second component defines the causality in terms particular system properties of interest to be computed from the model. The third component is the computation itself. Causality analysis can thus break down in various ways at each component. If the model does not adequately represent the system, the causality statement computed from the model may not pertain to the generative system. It may also be the case that the model structure does not easily afford computation of the intended definition of causality. Errors may also arise at the definitional level, for example if causality definition leads to logical inconsistencies. It is also possible that the definition of causality employed is not the actual property of interest to the investigator. Lastly, the inaccuracies in the method of computation may produce spurious results.

The major problem plaguing causality analyses in neuroscience is that the distinction between these separate components of the analysis, and possible points of failure, is not acknowledged. Instead, causality analyses are treated as a means of system identification—that with a particular causality statistic one may obtain an understanding of the system, without recognizing the model and assumptions that underly such causality statistics. Conversely, erroneous results of particular causality statistics when applied to particular generative systems are treated as failures of the particular method as a whole, as opposed to a failure in a particular component of the analysis process whose solution might ultimately be informative. The failure to distinguish between the various break-down points is emblematic of how the absence of such considerations and the ad hoc nature of causality analyses in neuroscience have resulted in serious ambiguity and high likelihood of misinterpretation.

In neuroscience investigations, particularly analysis of functional time series data, the ultimate objective is the identification, modeling, and analysis of the data gen-

erating system. This objective comprises a set of smaller more focused scientific questions, including localizing the components involved in the system, determining the directional relationships between components, ascertaining the directness of those connections, and characterizing the functional properties between components. Consequently, the causality framework of Pearl is inapplicable due to this uncertainty in the system under study, particularly at the whole brain level. Such methods may be appropriate in analysis of known neural networks, but in analysis of data from larger scale systems the system itself is unknown. For such analyses, the questions of interest pertain to various properties of the underlying system, but those properties can only be computed for the chosen model. Thus, there is persistent uncertainty in causality results as to whether the model corresponds to the system, in addition to whether the causality corresponds to the properties of interest. A question of interest may be whether causality methods are capable of identifying system properties even from simpler models, but answering this requires better understanding of the mapping from generative systems to the model class and then from the model to the causality. Ultimately, causality results can only be confirmed by experimental manipulation of the system.

Due to the confusion that clouds much of the causality analyses in neuroscience—the failure to acknowledge the different properties characterized by different methods and the failure to specifically diagnose points of break-down—it is my opinion that much of the focus on causality has been, not only superfluous, but detrimental to scientific investigations. Scientific questions of causality are really questions about specific properties of a chosen model, under the assumption that the model adequately represents the corresponding properties of the system. Therefore, much of the ambiguity can be resolved by clearly stating the properties of interest being computed without the burden of the philosophical quandary of causality. Because these properties express relationships between components under the model, it is also necessary to examine the dynamics of the individual components—nodes, channels, and inputs. Thus, my overall recommendation for causality analysis in neuroscience would be 1) to state the causality question in terms of specific properties of the system and 2)

answer that question using the broader modeling and system identification framework outlined in the next section.

6.5 Modeling and System Identification Framework

The modeling and system identification framework, as articulated by George Box [11], is a formal approach to data analysis, which closely parallels the scientific method. Figure 6-1 shows a general diagram of the modeling and system identification framework. With a scientific question in mind, one designs an experiment and gathers data. The same considerations that guided the experiment, shape how one models the data. They clarify the objective of the model, which suggests the appropriate class of models to evaluate and the criteria by which one selects a particular model fit. The chosen model estimate must be validated, such as by posterior predictive checks. If the model is found inadequate, such as colored noise or poor replication of the data, then the data must be remodeled with alternative model choices. If the model estimate is found adequate, it can then be analyzed to understand any properties of interest, and specifically to answer the motivating scientific question. The results of these analyses can then suggest follow up questions and guide subsequent experiments.

Considering this framework, it is clear that causality analyses are then a specific type of analysis of the properties of a chosen model. The identification of the system structure is achieved through the process of modeling. The veracity of statements of causality depends on the applicability of the model, and the extent of such applicability can only be verified experimentally.

Unfortunately, as discussed above and as indicated in Figure 6-2, the use of causality analysis in neuroscience is often intended, mistakenly, as a means of system identification itself. The process of modeling and the choices and assumptions that go into it are unacknowledged, if not completely hidden within a software toolbox. Thus, potential problems, such as inadequacy of the estimated model or computational difficulties with the analysis methods, indicated in Figure 6-3, are likely to go unad-

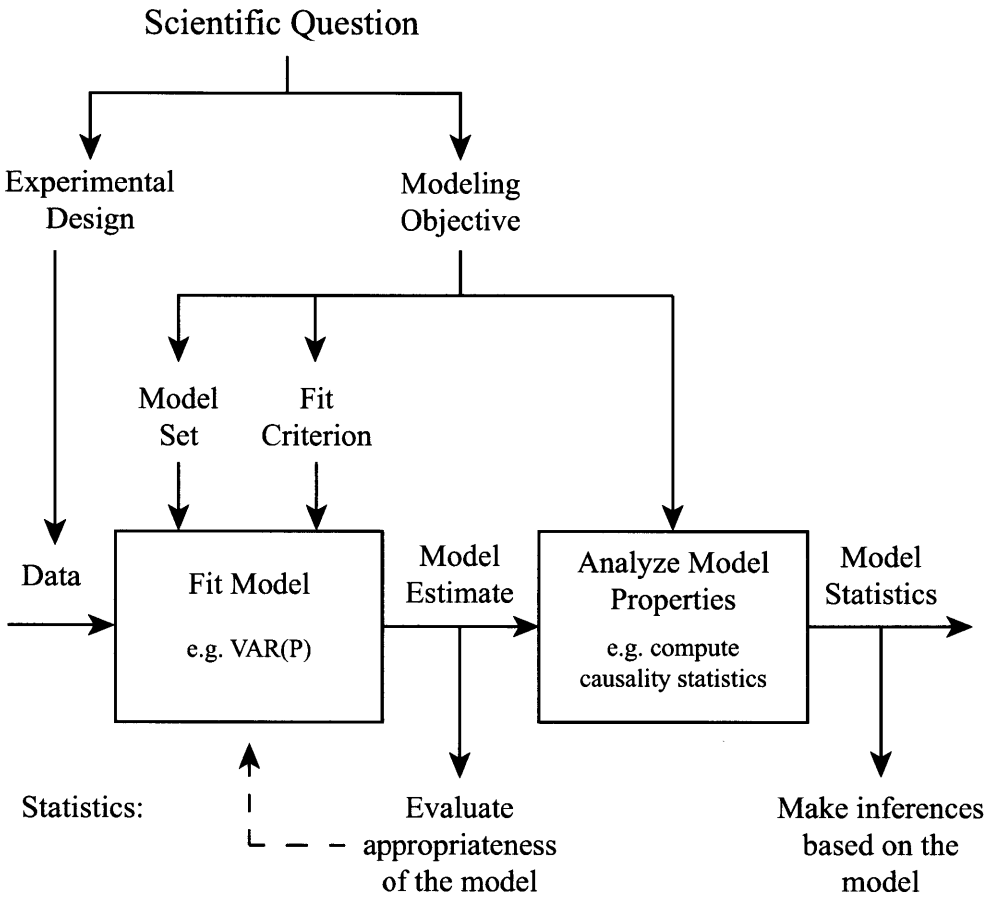

Figure 6-1: Modeling and System Identification Framework.

dressed. More fundamentally, the various causality measures are often conflated. The different causality measures can be more clearly characterized in terms of different scientific questions about system properties, and the misinterpretations of causality analyses undermine the investigation at more basic levels, as indicated in Figure 6-4. By understanding where causality analysis lies in the system identification framework, it is now possible to better pinpoint where causality methods breakdown.

As a concrete demonstration, we return to the epilepsy application in Section 1.2. Figure 1-1, reproduced from [20], is shown again below for convenience. Figure 6-5, shows causality estimates for two retrospective patients. The top plot shows the iEEG time series for retrospective patient 1. The middle two plots show the frequency-domain GG-causality between two electrodes (spH5 and dAH1) for the first and second halves of the time series. The bottom plot shows the iEEG time series for retrospective patient 3 with GG-causality estimates overlaid and an inset showing location of electrode placement. In both cases, the earliest significant causality occurs prior to and in different electrodes from the onset suggested by visual analysis of the iEEG. The interpretation of these results is that larger causality values indicate stronger influence or, equivalently, greater relative importance in determining the resulting system behavior, i.e. oscillations characteristic of seizure, and that GG-causality may be useful for planning surgical intervention. This interpretation is quite prevalent in neuroscience for this and other causality measures. However, GG-causality does not necessarily reflect the properties of interest in this case, and following the suggested modeling and system identification framework may be more informative.

The property of interest is the onset and spread of characteristic oscillations, for example large amplitude, high frequency oscillations, typical of seizure onset. Thus, one must select a model set, fit criterion, and estimation method that adequately models the oscillatory features of the data. It may be the case that linear time series models such as the VAR are appropriate. Any properties of the selected model can then be analyzed. The best understanding of the dynamics of the model comes from analyzing and plotting the internal dynamics of each node and the channels connecting them, as was done for the presentation of the systems for Examples 1 and 2 in

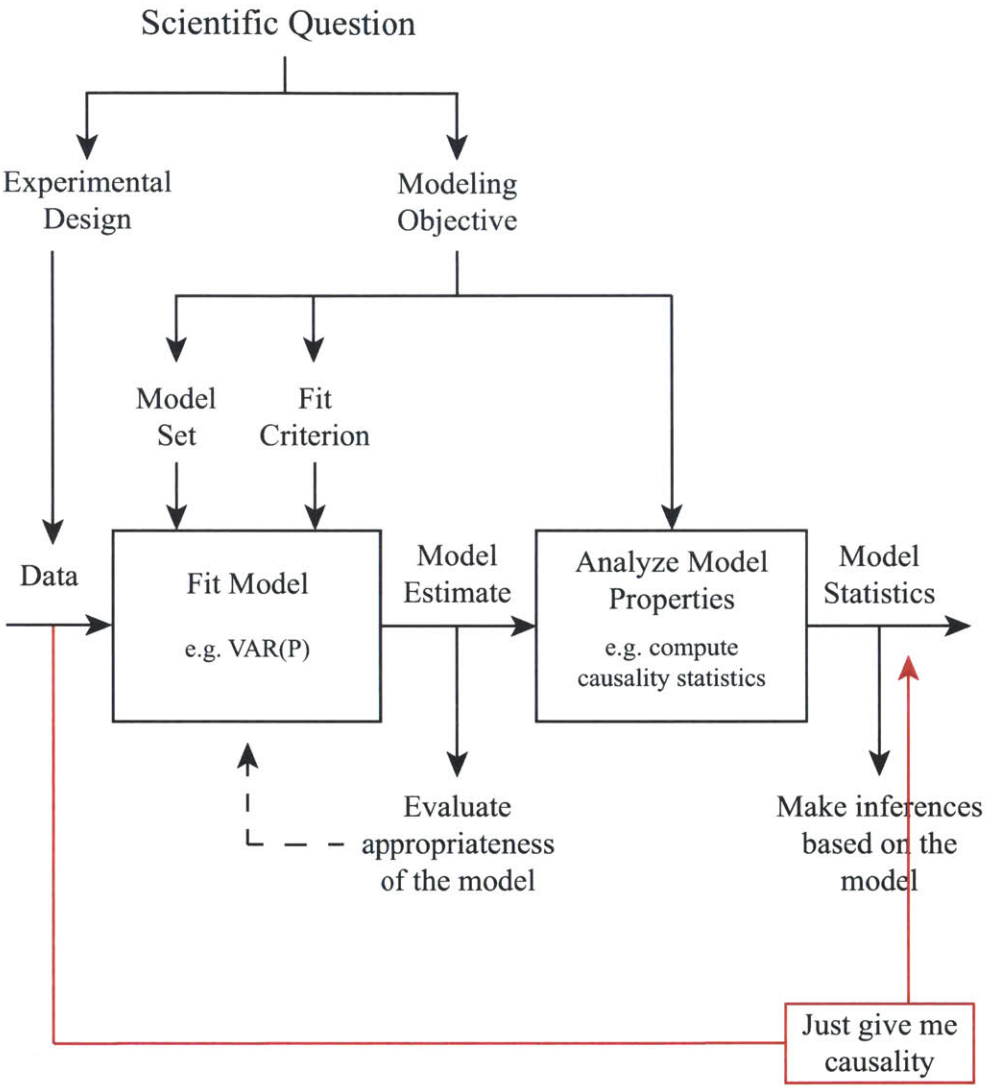

Figure 6-2: Modeling and System Identification Framework: How causality is often approached in neuroscience.

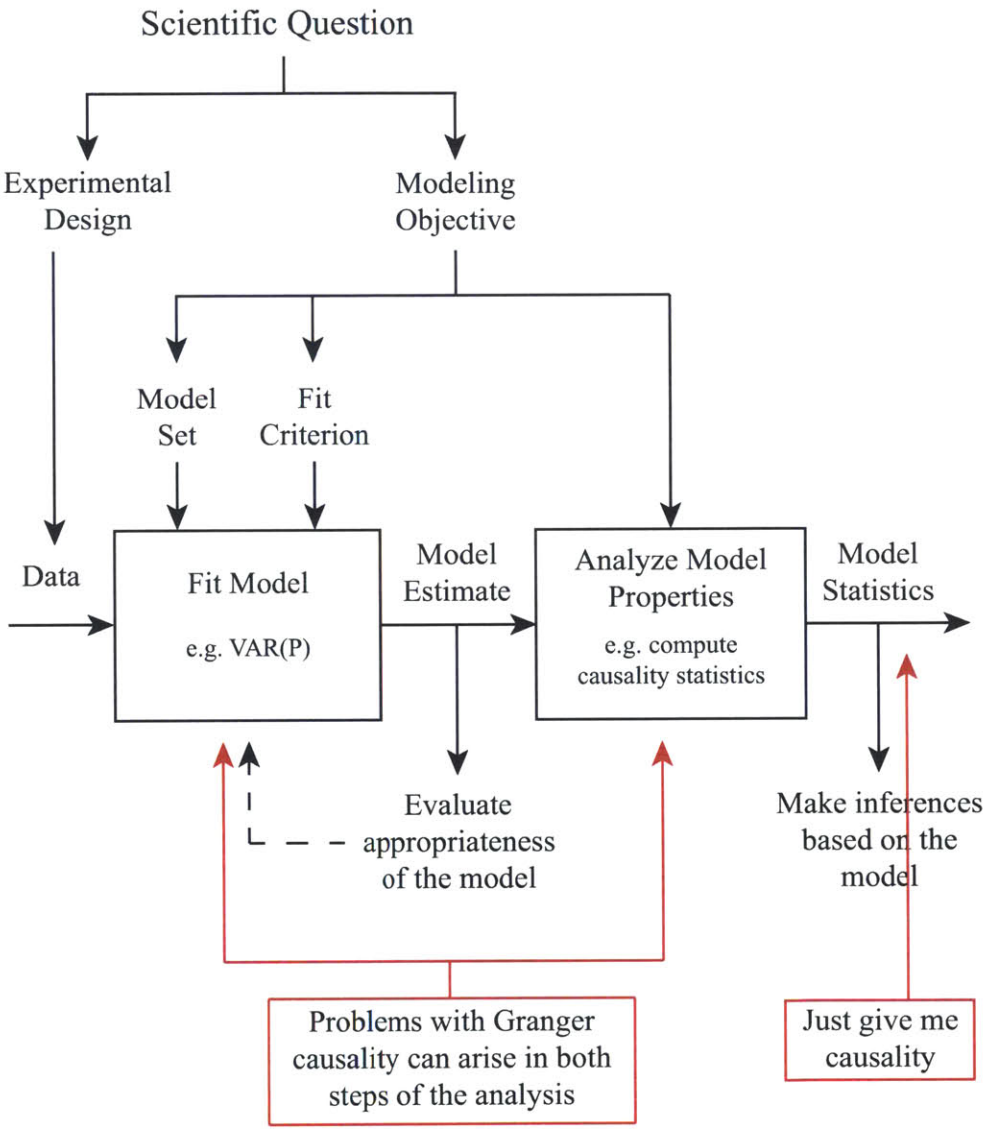

Figure 6-3: Modeling and System Identification Framework: Where problems arise in Granger causality.

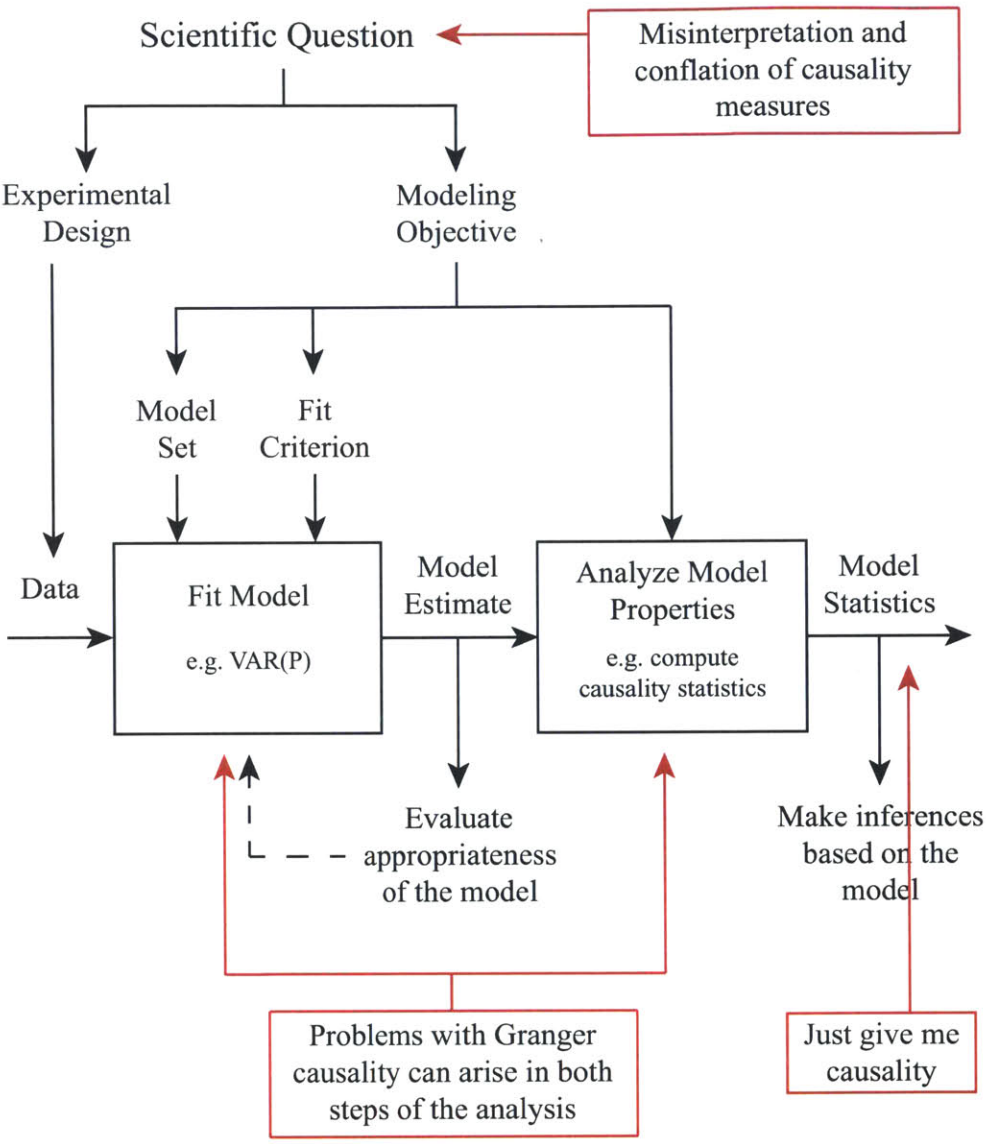

Figure 6-4: Modeling and System Identification Framework: A choice of causality determines the scientific question.

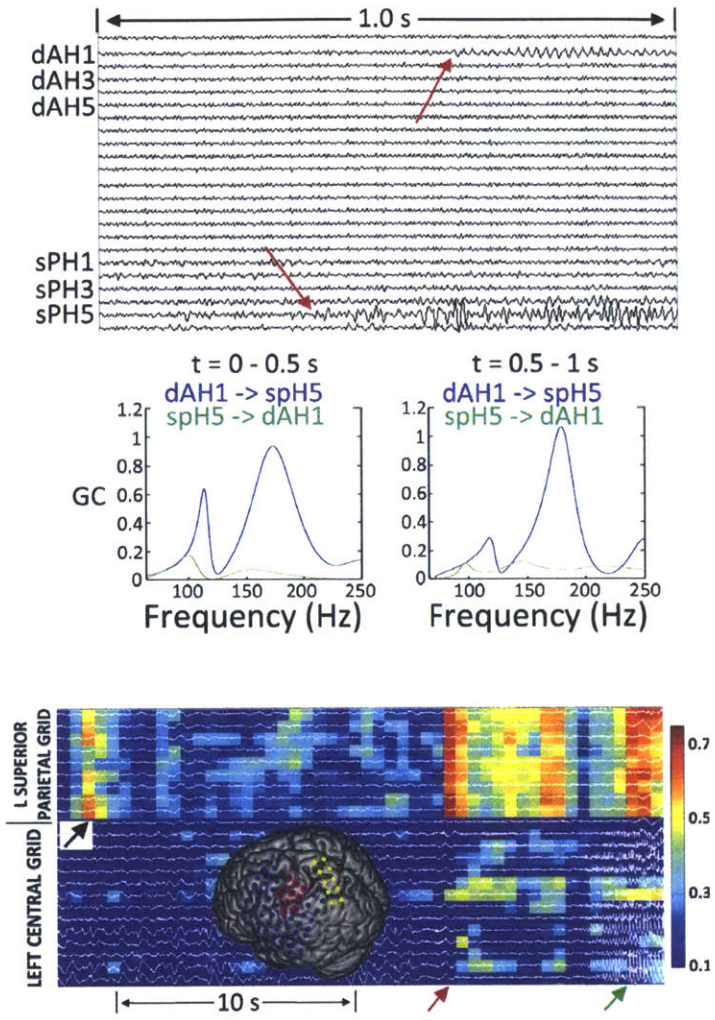

Figure 6-5: Causality estimates for two epilepsy patients. Reprinted from [20] by permission of publisher John Wiley & Sons, Inc.

Figures 3-1 and 4-1, respectively. In particular, it is likely informative to evaluate direct connections between components, such as via the 1-step detection test of Lutkepohl. It is then possible to analyze the specific properties of interest—in this case, the oscillations indicative of seizure in epilepsy. One can explore the source of large amplitude oscillations by analyzing the changes in component dynamics surrounding the time of seizure onset (i.e. Do the changes occur in the transmitter, receiver, or channel?). Additionally, because a further objective of such studies is improved planning of surgical intervention, it may also be informative to use the model to simulate resections by removing select nodes or channels from the model to investigate what actions most significantly prevent or reduce the oscillatory behavior characteristic of seizure. If other properties are of interest, such as improvement of prediction as quantified by GG-causality, then they can also be computed and deciphered in terms of the component dynamics. However, as suggested, improvement of prediction is likely not reflective of the properties of interest in the analysis of epilepsy data.

6.6 Conclusion

In this thesis, I have shown that there are fundamental flaws in use of Granger causality analysis in neuroscience:

I I have shown that computational problems arise in the current methods of estimating conditional GG-causality. The use of separate full and reduced model fits results in a peculiar bias-variance trade-off in model order and a high sensitivity to uncertainties in the estimated model parameters, leading to erroneous causality estimates, including spurious frequency-domain peaks and valleys and even negative values. The uncertainties introduced by these computational problems have thus far masked more fundamental interpretational problems that also undermine GG-causality analysis.

II I have formulated a method to correctly compute conditional GG-causality from a single full model estimate. With this correct GG-causality one can better

analyze the interpretational problems and better assess the utility of and one's interest in GG-causality analysis.

III I have shown that even with the correct GG-causality, and even in the simplest case of a bivariate, unidirectional system, GG-causality is not interpretable in terms of the system properties without evaluating the component dynamics of the estimated model. And more importantly, the propensity to misinterpret GG-causality in neuroscience is because the statement of GG-causality—improvement of prediction—is very likely not the notion of causality sought by neuroscientists.

There are many other causality measures and methods, each characterizing a different property of the system model, and causality analyses can produce erroneous results for various reasons: 1) the methods of computing/estimating causality may be flawed; 2) the data generating system may be insufficiently represented by the model; 3) the dynamic system properties of interest to the investigator may not be reflected by the causality measure. Indeed, each of these scenarios has presented itself neuroscience causality analyses. Unfortunately, there is much confusion in the literature due to conflation of the different causality measures and inarticulation of the specific points of failure. Much of this confusion can be resolved by appropriately placing causality analyses in a modeling and system identification framework, where any causality statement is seen to be the result of a scientific question about the properties of a chosen model of the data generating system.

Causality is concept which is both tantalizingly intuitive and deceptively elusive. Universal definitions and methods are still the subject of active philosophical investigations. While such tools could one day prove useful for systems analysis, they would likely be unnecessary for most neuroscience applications, which would more directly and clearly be addressed using the modeling and system identification framework.

Bibliography

[1] Brian D. O. Anderson and John B. Moore. Optimal Filtering. Prentice-Hall, Englewood Cliffs, N.J., 1979.

[2] Kristopher L. Anderson, Rajasimhan Rajagovindan, Georges A. Ghacibeh, Kimford J. Meador, and Mingzhou Ding. Theta oscillations mediate interaction between prefrontal cortex and medial temporal lobe in human memory. Cerebral Cortex, 20(7):1604–1612, July 2010.

[3] Laura Astolfi, Febo Cincotti, Donatella Mattia, M. Grazia Marciani, Luiz A. Baccala, Fabrizio De Vico Fallani, Serenella Salinari, Mauro Ursino, Melissa Zavaglia, Lei Ding, J Christopher Edgar, Gregory A. Miller, Bin He, and Fabio Babiloni. Comparison of different cortical connectivity estimators for high-resolution EEG recordings. Human Brain Mapping, 28(2):143–157, Feb 2007.

[4] L. A. Baccalá, K. Sameshima, G. Ballester, A. C. do Valle, and C. Timo-Iaria. Studying the interaction between brain structures via directed coherence and Granger causality. Applied Signal Processing, 5(1):40–48, 1998.

[5] Luiz A. Baccalá and Koichi Sameshima. Partial directed coherence: A new concept in neural structure determination. Biological Cybernetics, 84(6):463–474, Jan 2001.

[6] Lionel Barnett, Adam B. Barrett, and Anil K. Seth. Granger causality and transfer entropy are equivalent for Gaussian variables. Physical Review Letters, 103(23):238701, Dec 2009.

[7] Lionel Barnett and Anil K. Seth. The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference. Journal of Neuroscience Methods, 223:50–68, Feb 2014.

[8] Adam B. Barrett, Michael Murphy, Marie-Aurélie Bruno, Quentin Noirhomme, Mélanie Boly, Steven Laureys, and Anil K. Seth. Granger causality analysis of steady-state electroencephalographic signals during propofol-induced anaesthesia. Plos ONE, 7(1):e29072, January 2012.

[9] R. Clifford Blair and W. Karniski. An alternative method for significance testing of waveform difference potentials. Psychophysiology, 30(5):518–524, Sep 1993.

[10] Anil Bollimunta, Jue Mo, Charles E. Schroeder, and Mingzhou Ding. Neuronal mechanisms and attentional modulation of corticothalamic alpha oscillations. The Journal of Neuroscience, 31(13):4935–4943, March 2011.

[11] George E. P. Box, Gwilym M. Jenkins, and Gregory C. Reinsel. Time Series Analysis: Forecasting and Control. John Wiley, Hoboken, N.J., 2008.

[12] Steven L. Bressler and Anil K. Seth. Wiener–Granger causality: A well established methodology. NeuroImage, 58(2):323–329, September 2011.

[13] Andrea Brovelli, Mingzhou Ding, Anders Ledberg, Yonghong Chen, Richard Nakamura, and Steven L. Bressler. Beta oscillations in a large-scale sensorimotor cortical network: Directional influences revealed by Granger causality. Proceedings of the National Academy of Sciences, 101(26):9849–9854, Jun 2004.

[14] Alex J. Cadotte, Thomas B. Demarse, Thomas H. Mareci, Mansi B. Parekh, Sachin S. Talathi, Dong-Uk Hwang, William L. Ditto, Mingzhou Ding, and Paul R. Carney. Granger causality relationships between local field potentials in an animal model of temporal lobe epilepsy. Journal of Neuroscience Methods, 189(1):121–129, Jan 2010.

[15] Ryan T. Canolty and Robert T. Knight. The functional role of cross-frequency coupling. Trends in Cognitive Sciences, 14(11):506–515, Nov 2010.

[16] Yonghong Chen, Steven L. Bressler, and Mingzhou Ding. Frequency decomposition of conditional Granger causality and application to multivariate neural field potential data. Journal of Neuroscience Methods, 150(2):228–237, Jan 2006.

[17] J. Daunizeau, O. David, and K. E. Stephan. Dynamic causal modelling: A critical review of the biophysical and statistical foundations. NeuroImage, 58(2):312–322, Sep 2011.

[18] Jean-Marie Dufour and Eric Renault. Short run and long run causality in time series: Theory. Econometrica, 66(5):1099–1125, September 1998.

[19] Michael Eichler. A graphical approach for evaluating effective connectivity in neural systems. Philosophical Transactions of the Royal Society B, 360(1457):953–967, Jan 2005.

[20] C. M. Epstein, B. M. Adhikari, R. Gross, J. Willie, and M. Dhamala. Application of high-frequency Granger causality to analysis of epileptic seizures and surgical decision making. Epilepsia, 55(12):2038–2047, Dec 2014.

[21] Esther Florin, Joachim Gross, Johannes Pfeifer, Gereon R. Fink, and Lars Timmermann. Reliability of multivariate causality measures for neural data. Journal of Neuroscience Methods, 198(2):344–358, Jun 2011.

[22] Karl Friston, Jean Daunizeau, and Klaas Enno Stephan. Model selection and gobbledygook: Response to Lohmann et al. NeuroImage, 75:275–278; discussion 279–281, Jul 2013.

[23] Karl J. Friston, L. Harrison, and W. Penny. Dynamic causal modelling. NeuroImage, 19(4):1273–1302, Jan 2003.

[24] A. Gelman and J. Hill. Data Analysis Using Regression and Multilevel/Hierarchical Models. Analytical Methods for Social Research. Cambridge University Press, 2007.

[25] Felipe Gerhard, Tilman Kispersky, Gabrielle J. Gutierrez, Eve Marder, Mark Kramer, and Uri Eden. Successful reconstruction of a physiological circuit with known connectivity from spiking activity alone. PLoS Computational Biology, 9(7):e1003138, July 2013.

[26] John Geweke. Measurement of linear dependence and feedback between multiple time series. Journal of the American Statistical Association, 77(378):304–313, June 1982.

[27] John F. Geweke. Measures of conditional linear dependence and feedback between time series. Journal of the American Statistical Association, 79(388):907–915, December 1984.

[28] Steven Graham, Elaine Phua, Chun Siong Soon, Tomasina Oh, Chris Au, Borys Shuter, Shih-Chang Wang, and Ing Berne Ye. Role of medial cortical, hippocampal and striatal interactions during cognitive set-shifting. NeuroImage, 45(4):1359–1367, Jan 2009.

[29] Clive W. J. Granger. Investigation causal relations by econometric models and cross-spectral methods. Econometrica, 37(3):424–438, Jan 1969.

[30] Shuixia Guo, Anil K. Seth, Keith M. Kendrick, Cong Zhou, and Jianfeng Feng. Partial Granger causality – eliminating exogenous inputs and latent variables. Journal of Neuroscience Methods, 172(1):79–93, Jul 2008.

[31] Sanqing Hu and Hualou Liang. Causality analysis of neural connectivity: New tool and limitations of spectral Granger causality. Neurocomputing, 76(1):44–47, 2012.

[32] Ole Jensen and Laura L. Colgin. Cross-frequency coupling between neuronal oscillations. Trends in Cognitive Sciences, 11(7):267–269, Jul 2007.

[33] T. Kailath, A.H. Sayed, and B. Hassibi. Linear Estimation. Prentice-Hall information and system sciences series. Prentice Hall, 2000.

[34] M. J. Kamiński and Katarzyna J. Blinowska. A new method of the description of the information flow in the brain structures. Biological Cybernetics, 65(3):203–210, Jan 1991.

[35] Maciej Kamiński, Mingzhou Ding, Wilson A. Truccolo, and Steven L. Bressler. Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. Biological Cybernetics, 85(2):145–157, Jan 2001.

[36] Sanggyun Kim, David Putrino, Soumya Ghosh, and Emery N. Brown. A Granger causality measure for point process models of ensemble neural spiking activity. PLoS Computational Biology, 7(3):e1001110, March 2011.

[37] Tilman Kispersky, Gabrielle J. Gutierrez, and Eve Marder. Functional connectivity in a rhythmic inhibitory circuit using Granger causality. Neural Systems & Circuits, 1(9):1–15, may 2011.

[38] Anna Korzeniewska, Ciprian M. Crainiceanu, Rafal Kus, Piotr J. Franaszczuk, and Nathan E. Crone. Dynamics of event-related causality in brain electrical activity. Human Brain Mapping, 29(10):1170–1192, Jan 2008.

[39] Anna Korzeniewska, Malgorzata Manczak, Maciej Kamiński, Katarzyna J. Blinowska, and Stefan Kasicki. Determination of information flow direction among brain structures by a modified directed transfer function (dDTF) method. Journal of Neuroscience Methods, 125(1-2):195–207, Jan 2003.

[40] Gabriele Lohmann, Kerstin Erfurth, Karsten Müller, and Robert Turner. Critical comments on dynamic causal modelling. NeuroImage, 59(3):2322–2329, Feb 2012.

[41] Gabriele Lohmann, Karsten Müller, and Robert Turner. Response to commentaries on our paper: Critical comments on dynamic causal modelling. NeuroImage, 75:279–281, 2013.

[42] Helmut Lütkepohl. New Introduction to Multiple Time Series Analysis. Springer, Berlin; New York, 2005.

[43] J. Roderick McCrorie and Marcus J. Chambers. Granger causality and the sampling of economic processes. Journal of Econometrics, 132(2):311–336, 2006.

[44] E. Olbrich and P. Achermann. Analysis of oscillatory patterns in the human sleep EEG using a novel detection algorithm. Journal of Sleep Research, 14(4):337–346, Jan 2005.

[45] Judea Pearl. Causality: Models, Reasoning, and Inference. Cambridge University Press, Cambridge, U.K.; New York, 2000.

[46] Christopher J. Quinn, Todd P. Coleman, Negar Kiyavash, and Nicholas G. Hatsopoulos. Estimating the directed information to infer causal relationships in ensemble neural spike train recordings. Journal of Computational Neuroscience, 30(1):17–44, Feb 2011.

[47] Alard Roebroeck, Elia Formisano, and Rainer Goebel. Mapping directed influence over the brain using Granger causality and fMRI. NeuroImage, 25(1):230–242, Jan 2005.

[48] Thomas Schreiber. Measuring information transfer. Physical Review Letters, 85(2):461–464, July 2000.

[49] Anil K Seth. A MATLAB toolbox for granger causal connectivity analysis. Journal of Neuroscience Methods, 186(2):262–273, Feb 2010.

[50] Anil K. Seth, Adam B. Barrett, and Lionel Barnett. Causal density and integrated information as measures of conscious level. Philosophical Transactions of the Royal Society A, 369(1952):3748–3767, Oct 2011.

[51] S Wang, Y Chen, M Ding, J Feng, J. F. Stein, T. Z. Aziz, and X. Liu. Revealing the dynamic causal interdependence between neural and muscular signals in Parkinsonian tremor. Journal of the Franklin Institute, 344(3):180–195, Jan 2007.

[52] Xue Wang, Yonghong Chen, Steven L. Bressler, and Mingzhou Ding. Granger causality between multiple interdependent neurobiological time series: Blockwise versus pairwise methods. International Journal of Neural Systems, 17(2):71–78, Apr 2007.