9-29-2010

ESTIMATING TEMPORAL ASSOCIATIONS IN ELECTROCORTICOGRAPHIC (ECoG) TIME SERIES WITH FIRST ORDER PRUNING

Haley Hedlin
Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics, haleyhedlin@gmail.com

Dana Boatman
Johns Hopkins University, School of Medicine, Neuro Epilepsy

Brian Caffo
Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics

Estimating temporal associations in electrocorticographic (ECoG) time series with first order pruning

Haley Hedlin, Dana Boatman, and Brian Caffo

September 23, 2010

Abstract

Granger causality (GC) is a statistical technique used to estimate temporal associations in multivariate time series. Many applications and extensions of GC have been proposed since its formulation by Granger in 1969. Here we control for potentially mediating or confounding associations between time series in the context of event-related electrocorticographic (ECoG) time series. A pruning approach to remove spurious connections and simultaneously reduce the required number of estimations to fit the effective connectivity graph is proposed. Additionally, we consider the potential of adjusted GC applied to independent components as a method to explore temporal relationships between underlying source signals. Both approaches overcome limitations encountered when estimating many parameters in multivariate time-series data, an increasingly common predicament in today's brain mapping studies.

1 Introduction

Within the brain, connectivity is categorized into three types, structural, functional, and effective, to describe the physical and electrical connections at various levels of brain anatomy

and function. Structural connectivity examines the anatomical structure of the connections, such as fiber tracts. Functional connectivity refers to "temporal correlations between remote neurophsyiological events." Effective connectivity refers to "the influence one neural system exerts over another" (Friston, 1994). This work focuses on estimating effective connectivity between populations of neurons. The approach taken allows many signals to be simultaneously considered while avoiding the identification of spurious connections between the signals.

Specifically, our work aims to estimate temporal relationships between multivariate time series measuring electrical field potentials generated by a population of neurons while adjusting for potentially mediating or confounding associations between time series. These measurements are most commonly recorded across the surface of the scalp as in electroencephalography (EEG). However, the recordings are occasionally measured from multichannel electrodes placed directly on the surface of the cortex in patients who undergo neurosurgery. These high dimensional recordings are referred to as electrocorticography (ECoG) or intracranial EEG (for an accessible and detailed description of the activity recorded by EEG and ECoG see Bressler and Ding (2002) and Boatman-Reich et al. (2010)). In both EEG and ECoG, each recording produces 20-300 simultaneous time series (Oostenveld and Praamstra, 2001). Common statistical methods to infer functional and effective connectivity using EEG or ECoG data include coherence (Walter, 1963; Shaw, 1981), dynamic causal modeling (David et al., 2006), Granger causality (Granger, 1969), and, closely related to Granger causality, directed transfer function (Kamiński et al., 2001; Blinowska et al., 2004). More generally, surveys of brain connectivity can be found in Bullmore and Sporns (2009); David et al. (2005); Friston (2004); Horwitz (2003).

In particular, we estimate temporal relationships between ECoG time series using Granger causality (GC). We apply GC as an exploratory technique, as opposed to confirmatory techniques such as dynamic causal modeling. Several works have recently included applications of GC to ECoG and EEG data; see Marinazzo et al. (2010); Dauwels et al. (2009) and Oya

et al. (2007). GC has been extensively applied to functional magnetic resonance imaging (fMRI) data in addition to EEG and ECoG (see Havlicek et al., 2010; Jiao et al., 2010; Liao et al., 2010; Sato et al., 2010, for recent applications to fMRI). Many variations of Granger's original formulation have been proposed in the context of EEG, ECoG, and fMRI, each extending the applicability of the method to specific data structures and research questions. Bressler and Seth (2010) provide an excellent review of Granger causality and Ding et al. (2006) delineate the mathematical frameworks of Granger causality.

The first applications of Granger causality to neuroscience data considered all pairwise connections between time series to generate directed graphs representing effective connectivity networks (Kamiński et al., 2001). However, such bivariate analyses cannot differentiate between direct connections ($\mathbf{X} \rightarrow \mathbf{Y}$) and connections mediated through a third time series ($\mathbf{X} \rightarrow \mathbf{Z} \rightarrow \mathbf{Y}$) or confounded by a third time series ($\mathbf{Z} \rightarrow \mathbf{X}$ and $\mathbf{Z} \rightarrow \mathbf{Y}$). Adjusted GC avoids misrepresenting mediated connections as direct by considering a multivariate autoregressive (MVAR) model that includes time series $\mathbf{Z}$ as an adjustor of the relationship between time series $\mathbf{X}$ and $\mathbf{Y}$. (Geweke, 1984; Ding et al., 2006). Adjusted GC also avoids identifying spurious connections arising from confounding between time series. Multiple adjustor time series may be included for data sets involving more than three time series. The number of time series we may include using this approach is limited, however, as estimates from MVAR models with many time series become increasingly unstable as the dimensions of the MVAR increase (Lütkepohl, 2005). This limitation is particularly troublesome as current brain mapping studies using EEG now routinely use high density multichannel recordings with up to 128 channels (Oostenveld and Praamstra, 2001; Lantz et al., 2003).

Because fitting a fully adjusted MVAR model to datasets with many time series is not feasible, we have devised a "pruning" approach to remove spurious bivariate connections using adjusted Granger causality. The pruning approach allows us to apply adjusted Granger causality to a larger number of time series than has been done in the past while avoiding fitting excessively large MVAR models, limiting the number of MVAR models being fit, and

reducing the number of spurious connections estimated in the resulting graphs representing effective connectivity. Here we consider only first-order pruning, i.e. fitting MVAR models that include a single adjustor time series when estimating GC. We demonstrate using real and simulated data that the pruning approach to adjusted Granger causality greatly reduces the number of spurious connections estimated in multichannel settings.

Another approach to reduce the dimensionality of the MVAR models first applies independent components analysis (ICA) to the data. ICA serves both as a data reduction step and as an approach to uncovering the networks underlying the raw data. The dimension reduction of the data permits a fully adjusted GC model to be fit to the retained components, i.e all components are included as adjustors in the model. The resulting GC graph describes the connectivity between the underlying networks represented by the independent components.

Here we consider methods for estimating adjusted Granger causality in the time domain of multiple simultaneous EEG or ECoG time series. The methods overcome limitations on the number of time series included in the analysis while concurrently considering adjustments to control for the possible confounding or mediating associations between neighboring or remote time series. We begin with a general formulation of Granger causality in Section 3. The method of pruning to fit large effective connectivity graphs via Granger causality is covered in Section 4, along with simulations and a real data application in Sections 5 and 6, respectively. The method of network Granger causality via ICA is described in Section 7 and we conclude with a discussion in Section 8.

2 Background

Granger proposed identifying the temporal orderings of simultaneous time series using the framework of autoregressive models (Granger, 1969). His method formalized concepts of causality conceived by Wiener (Wiener, 1956; Ding et al., 2006). Granger's approach ex-

tended the statistical concepts of coherence and partial correlation to directional causality, specifically linear dependence and feedback (Geweke, 1982). In short, $\mathbf{Y}$ is said to Granger cause $\mathbf{X}$ if adding $\mathbf{Y}$ as an adjustor to the autoregressive model of $\mathbf{X}$ reduces the variance of $\mathbf{X}$ from the level observed when only previous observations in $\mathbf{X}$ are considered.

Geweke later defined quantities to estimate the directional causality between two time series described in Granger's formulation (Geweke, 1982). These quantities have the desirable property of being equal to zero when dependence between the two time series is absent. Under this definition the dependence between two time series could be easily partitioned into three such quantities. In 1984, Geweke further extended these ideas to allow for the dependence between two time series to be conditional on a third time series (Geweke, 1984). Since its introduction, this concept has been called "conditional Granger causality" (Ding et al., 2006). With this extension it became possible to identify spurious connections that are solely mediated or confounded by a third time series.

In the past, applications of GC to neuroscience considered only pairwise relationships between time series and combined these relationships to understand the broader relationships. Recent work extends the pairwise relationships to consider a potentially mediating or confounding association with a third time series. Ding et al. (2006) provide an introduction to conditional Granger causality in the temporal and frequency domains. Several authors have proposed nonlinear extensions to conditional Granger causality recently (see Chen et al., 2004; Gao and Tian, 2009). A nonparametric approach to estimating spectral conditional Granger causality is proposed by Dhamala et al. (2008). The above works assume that all relevant variables are observed. To address the violation of this assumption, Guo et al. (2008) extends linear and nonlinear conditional Granger causality to include exogenous and latent variables. Finally, others have considered time-varying applications of Granger causality (see Arnold et al., 1998; Ding et al., 2000; Hesse et al., 2003; Sato et al., 2010).

Despite the well established literature applying Granger causality to neuroscience data, few applications involve more than a dozen time series. Of those applications that consider

more time series, the typical approach is to consider only pairwise relationships or to reduce dimensionality of the data by using regions of interest (ROIs) (e.g. Hesse et al., 2003; Astolfi et al., 2007; Marinazzo et al., 2010). The aforementioned approaches all theoretically extend to larger numbers of time series, but in practice the estimates from MVAR models with many time series are poorly estimated (Lütkepohl, 2005). As the number of time series per individual continues to rise in brain mapping research, a practical approach to obtaining accurate estimates of graphs describing effective connectivity will become increasingly vital.

3 Formulation of Granger causality

3.1 Bivariate Granger causality

Consider two wide-sense stationary time series, $\mathbf{X}$ and $\mathbf{Y}$. We say $\mathbf{Y}$ G-causes $\mathbf{X}$ if the variance in the autoregressive model on $\mathbf{X}$ is greater than when $\mathbf{Y}$ is included as a predictor. Formally, in the notation of Geweke (1982, 1984) and Ding et al. (2006), the individual autoregressive models are

$$\begin{aligned} X_t &= \sum_{j=1}^{\infty} a_{1j} X_{t-j} + \epsilon_{1t}, \quad \mathrm{Var}(\epsilon_{1t}) = \Sigma_1 \\ Y_t &= \sum_{j=1}^{\infty} q_{1j} Y_{t-j} + \gamma_{1t}, \quad \mathrm{Var}(\gamma_{1t}) = \Gamma_1 \end{aligned} \tag{1}$$

and, the joint MVAR model

$$\begin{aligned} X_t &= \textstyle\sum_{j=1}^{\infty} a_{2j} X_{t-j} + \sum_{j=1}^{\infty} b_{2j} Y_{t-j} + \epsilon_{2t} \\ Y_t &= \textstyle\sum_{j=1}^{\infty} p_{2j} X_{t-j} + \sum_{j=1}^{\infty} q_{2j} Y_{t-j} + \gamma_{2t} \end{aligned} \qquad \mathbf{\Omega}_2 = \begin{pmatrix} \Sigma_2 & \Psi_2 \\ \Psi_2 & \Gamma_2 \end{pmatrix} \tag{2}$$

where $\mathrm{Cov}(\epsilon_{2t}, \gamma_{2t}) = \mathbf{\Omega}_2$ is the covariance matrix of the joint autoregression. We assume that the models are zero mean, without loss of generality.

$\mathbf{Y}$ is said to Granger cause $\mathbf{X}$ if $\text{Var}(\epsilon_{2t}) < \text{Var}(\epsilon_{1t})$. Geweke introduced the measure

$$F_{Y \to X} = \ln \frac{\Sigma_1}{\Sigma_2} \tag{3}$$

to quantify the Granger causal influence of $\mathbf{X}$ on $\mathbf{Y}$. We will denote $\mathbf{Y}$ Granger causes $\mathbf{X}$ by $\mathbf{Y} \to \mathbf{X}$. Similarly, $\mathbf{X} \to \mathbf{Y}$ if $\text{Var}(\gamma_{2t}) < \text{Var}(\gamma_{1t})$ and $F_{X \to Y} = \ln \frac{\Gamma_1}{\Gamma_2}$. Geweke also defined the interdependence of $\mathbf{X}$ and $\mathbf{Y}$ (sometimes called "instantaneous causality", see Ding et al., 2006; Geweke, 1984) as $F_{X \cdot Y} = \ln \frac{\Sigma_2 \Gamma_2}{|\mathbf{\Omega_2}|}$. Hence, the total dependence between $\mathbf{X}$ and $\mathbf{Y}$ can be written as $F_{X,Y} = \ln \frac{\Sigma_1 \Gamma_1}{|\mathbf{\Omega_2}|} = F_{X \to Y} + F_{Y \to X} + F_{X \cdot Y}$ (Geweke, 1982, 1984).

To formalize the directed graphs we will be using to describe the connectivity relationships estimated via GC, each node represents a time series recorded from an electrode and a directed edge $\mathbf{X} \to \mathbf{Y}$ denotes the temporal precedence of $\mathbf{X}$ on $\mathbf{Y}$ as estimated by Granger causality, where both $\mathbf{X}$ and $\mathbf{Y}$ are time series. We will refer generally to the temporal relationship between $\mathbf{X}$ and $\mathbf{Y}$ as a connection throughout. Note that $X \to X$ is not of interest, so it is omitted from all graphs and is assumed to be implied. More information on graphical representation is given in Section 3.3.

3.2 Adjusted Granger causality

In the previous section we only considered the relationship between two time series (called pairwise or bivariate GC). Consider the possibility that the Granger causal relationship $\mathbf{Y} \to \mathbf{X}$ is mediated or confounded by one or more time series $\mathbf{Z}$. We say $\mathbf{Y}$ G-causes or temporally precedes $\mathbf{X}$ when adjusted for $\mathbf{Z}$ ($\mathbf{Y} \to \mathbf{X}|\mathbf{Z}$) if the estimated GC coefficient

$$F_{Y \to X|Z} = \ln \frac{\Sigma_3}{\Sigma_4} \tag{4}$$

exceeds some prespecified threshold (discussed in Section 4.1.2) where Σ_3 is the variance of the autoregressive process of $\mathbf{X}$ when adjusting for $\mathbf{Z}$ alone and Σ_4 is the variance of $\mathbf{X}$ in the MVAR joint with $\mathbf{Y}$ and $\mathbf{Z}$ (as described in Equations (5) and (6) below).

$$
\begin{aligned}
X_t &= \textstyle\sum_{j=1}^{\infty} a_{3j} X_{t-j} + \sum_{j=1}^{\infty} c_{3j} Z_{t-j} + \epsilon_{3t} \\
Z_t &= \textstyle\sum_{j=1}^{\infty} u_{3j} X_{t-j} + \sum_{j=1}^{\infty} w_{3j} Z_{t-j} + \eta_{3t}
\end{aligned}
\qquad
\boldsymbol{\Omega}_3 = \begin{pmatrix} \Sigma_3 & \Psi_3 \\ \Psi_3 & \Upsilon_3 \end{pmatrix} \tag{5}
$$

$$
\begin{aligned}
X_t &= \textstyle\sum_{j=1}^{\infty} a_{4j} X_{t-j} + \sum_{j=1}^{\infty} b_{4j} Y_{t-j} + \sum_{j=1}^{\infty} c_{4j} Z_{t-j} + \epsilon_{4t} \\
Y_t &= \textstyle\sum_{j=1}^{\infty} p_{4j} X_{t-j} + \sum_{j=1}^{\infty} q_{4j} Y_{t-j} + \sum_{j=1}^{\infty} r_{4j} Z_{t-j} + \gamma_{4t} \\
Z_t &= \textstyle\sum_{j=1}^{\infty} u_{4j} X_{t-j} + \sum_{j=1}^{\infty} v_{4j} Y_{t-j} + \sum_{j=1}^{\infty} w_{4j} Z_{t-j} + \eta_{4t}
\end{aligned}
\tag{6}
$$

$$
\boldsymbol{\Omega}_4 = \begin{pmatrix} \Sigma_4 & \Psi_{4,1} & \Psi_{4,2} \\ \Psi_{4,1} & \Gamma_4 & \Psi_{4,3} \\ \Psi_{4,2} & \Psi_{4,3} & \Upsilon_4 \end{pmatrix}
$$

3.3 Graphical representation

We will use graphs to represent the estimated GC relationships. Each time series is represented by a node in the graph. Connections between the nodes are represented by directed edges. We refer to a graph created from estimating the bivariate connections as a bivariate graph. An edge from $\mathbf{X}$ to $\mathbf{Y}$ is determined to be present in the bivariate graph if the GC coefficient $F_{X \to Y}$ exceeds a predetermined threshold. In an adjusted graph, an edge from $\mathbf{X}$ to $\mathbf{Y}$ is present if the GC coefficient $F_{X \to Y|Z}$ exceeds a predetermined a threshold for all observed time series $\mathbf{Z}$. Hereafter, when $\mathbf{Z}$ refers to all observed time series, the resulting graph will be referred to as a fully-adjusted graph. If $\mathbf{Z}$ is a single time series, then the resulting graph is a first-order pruned adjusted graph. Finally, note that the threshold used to create the adjusted graph is possibly a different threshold than the threshold used to create the bivariate graph. For information on graphical models see Lauritzen (1996) and for more on the graphical representation of GC relationships, see Eichler (2005)

4 Pruning approach to adjusted GC

To estimate temporal relationships between large numbers of time series using adjusted GC, we take a pruning approach. Such an approach allows simultaneous consideration of many time series while removing spurious connections estimated by pairwise GC. The method involves fitting only low dimensional MVAR models; however, fitting many low dimensional models to fully describe the connections between many time series could quickly build to an undesirable level. With K time series, considering all pairwise connections using GC requires fitting $3 * \begin{pmatrix} K \\ 2 \end{pmatrix}$ bivariate autoregression models. That is, for each pair of time series, we must fit the two independent models and the joint model. The number of models to be estimated can be reduced to $K + \begin{pmatrix} K \\ 2 \end{pmatrix}$ simply by saving covariance estimates to be used in multiple pairs' GC estimates. To fit adjusted trivariate autoregressive models with each time series as a possible adjustor between the pairwise connections requires fitting $(K-1) * \begin{pmatrix} K \\ 2 \end{pmatrix}$ trivariate autoregressive models. Even with modest increases in K, the number of models to be fit increases rapidly. Fortunately, not all trivariate MVAR models will need to be fit in the pruning approach.

Spurious connections can appear in the presence of mediating or confounding associations between time series. Hence, we use the estimated bivariate GC graph to identify potential mediator or confounder time series. A mediator is one or more time series $\mathbf{Z}$ such that $\mathbf{X} \to \mathbf{Z} \to \mathbf{Y}$ and $\mathbf{X} \not\to \mathbf{Y}$. See Figure 5.1 for an example. A confounder, on the other hand is a time series $\mathbf{Z}$ such that $\mathbf{Z} \to \mathbf{X}$ and $\mathbf{Z} \to \mathbf{Y}$ but $\mathbf{X} \not\to \mathbf{Y}$ nor vice versa. This approach is most useful when the bivariate GC graph is relatively sparse, i.e. the number of pairwise connections is limited. Each bivariate connection has only a subset of the time series that are possible mediators or confounders. These possible mediators and confounders will be included as adjustors when fitting the multivariate GC model. As a result, instead of fitting all possible trivariate models, we fit only the models with adjustors from the subset.

To estimate the adjusted GC graph representing the effective connectivity in the multivariate time series, we first examine all possible bivariate pairs of the time series. The bivariate GC is estimated for each pair of time series using the variance components estimated in Equations (1) and (2). All bivariate GC estimates that exceed a prespecified threshold contribute an edge to the bivariate GC graph. Next, the bivariate GC graph is pruned to remove spurious edges. The spurious edges we consider are due to a third time series that either mediates or confounds the relationship between the time series connected by the edge. The pruning is accomplished by fitting the adjusted model described in Equations (5) and (6) to only those connection and mediator/confounder pairs that are possible given the bivariate GC graph. If an edge is determined to be due solely to mediating and/or confounding associations between time series, then it is considered to be a spurious connection and the edge is removed. The threshold applied to the adjusted model may differ from the threshold applied to the binary model. The remaining connections form our final mediator and confounder adjusted GC graph.

4.1 Algorithm

Algorithm for estimating mediator and confounder adjusted GC:

1. Choose thresholds that account for multiple comparisons and a model order that minimizes the criterion of choice (BIC, AIC, FPE, etc.) of the fitted MVAR model

2. Estimate bivariate GC coefficients for all pairs of time series in both directions

3. Create a 0/1 matrix R indicating which bivariate coefficients exceed the threshold (note: dim(R)=nchan x nchan and diag(R) equals 0 because we are not concerned with $\mathbf{X} \rightarrow \mathbf{X}$) where the rows of R correspond to the preceding time series and the columns of R correspond to the "postceding" time series

4. For each each bivariate GC coefficient (say $F_{X \rightarrow Y}$) exceeding the prespecified threshold, identify candidate mediating and confounding time series ($\mathbf{Z}$)

- Identify possible mediating time series *
- Identify possible confounding time series **

5. Estimate adjusted GC coefficients from trivariate models for each connection and adjustor pair (i.e. $\mathbf{X} \rightarrow \mathbf{Y}$ and $\mathbf{Z}$ sets) identified in the previous step

6. Identify which adjusted connections exceed the threshold

7. If the adjusted connections do not exceed the threshold for any one of the adjusted GC coefficients estimated in Step 5, then the bivariate connection is considered spurious, i.e. the connection solely exists through a mediating or confounding time series, and the edge is removed from the estimated graph

* Any time series $\mathbf{Z}$ that "postcedes" $\mathbf{X}$ (i.e. $\mathbf{X} \rightarrow \mathbf{Z}$) and precedes $\mathbf{Y}$ ($\mathbf{Z} \rightarrow \mathbf{Y}$) is a possible mediator of the $\mathbf{X} \rightarrow \mathbf{Y}$ connection. Such candidate mediator time series $\mathbf{Z}$ can be identified from the 0/1 matrix created from the bivariate GC estimates (R). Consider the row of R that corresponds to $\mathbf{X}$ (call it $R_{X,\bullet}$) and the column of R that corresponds to $\mathbf{Y}$ ($R_{\bullet,Y}$). Any elements that are equal to 1 in both $R_{X,\bullet}$ and $R^T_{\bullet,Y}$ indicate a possible mediator $\mathbf{Z}$ (where T denotes matrix transpose).

** Any time series $\mathbf{Z}$ that precedes $\mathbf{X}$ and precedes $\mathbf{Y}$ is a possible confounder in the relationship between $\mathbf{X}$ and $\mathbf{Y}$. Such candidate confounder time series can be identified in a manner similar to the candidate mediators in *. Any element that equals 1 in both $R_{\bullet,X}$ and $R_{\bullet,Y}$ is a possible confounder of connections between $\mathbf{X}$ and $\mathbf{Y}$.

4.1.1 Example

Consider the following example 0/1 matrix describing hypothetical bivariate GC connections in a set of 5 time series. Element j, k equal to 1 means that the time series in the j^{th} row Granger causes the time series in the k^{th} column. A 0 means no such connection exists. Labels A through E are given to each time series to aid discussion.

	A	B	C	D	E
A	0	0	0	0	0
B	1	0	1	0	0
C	0	1	0	1	1
D	0	1	0	0	1
E	1	1	1	0	0

Consider the $E \rightarrow A$ connection (element (5,1) in the matrix). To identify possible mediators of this connection, we examine row $R_{E,\bullet}$ (the fifth row, [1 1 1 0 0]) and column $R_{\bullet,A}$ (the first column, [0 1 0 0 1]) for elements that are equal to 1 in both vectors. The only element that equals 1 in both vectors is the second element, corresponding to B. Therefore, B is a candidate mediator of the $E \rightarrow A$ connection and we will adjust for B when estimating the GC for $E \rightarrow A$ to determine if the connection is solely mediated by B or if there is a direct connection. A direct connection can exist in two situations: in addition to mediation by B or alone without any mediation by B).

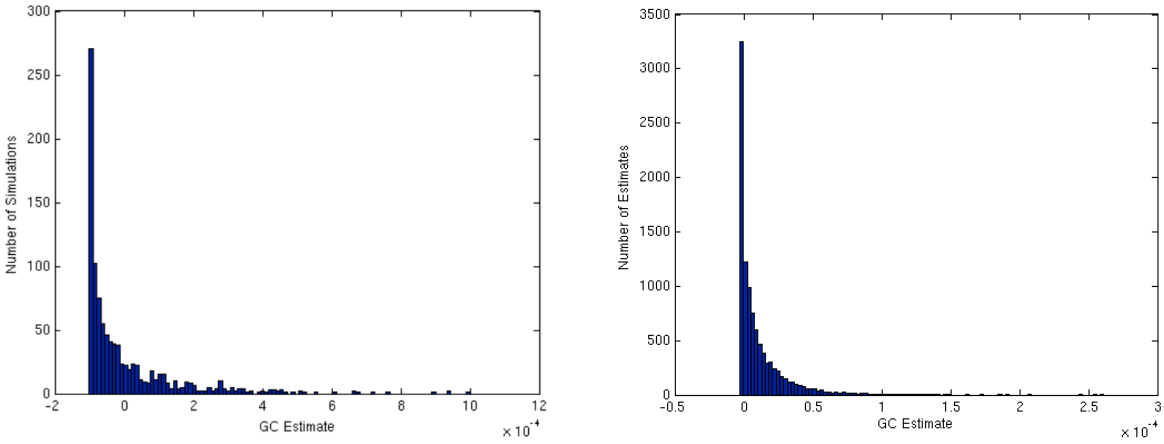
Next consider the $B \rightarrow C$ connection. To identify possible confounders, we examine columns $R_{\bullet,B}$ and $R_{\bullet,C}$ (i.e. the second and third columns). The fifth element is the only element to equal 1 in both columns, so E is identified as the only candidate confounder of the $B \rightarrow C$ connection. As a result, when estimating the adjusted GC, we will adjust for E when estimating the GC for $B \rightarrow C$.

4.1.2 Threshold choice

In the simulated data, the threshold is chosen by running 500 triples of independent autoregressive time series of length 10,000 (after discarding the first 1,000 as burn-in) with coefficients and covariances of magnitudes similar to those in the simulations presented in Section 5. The adjusted GC coefficient defined in Equation (4) is estimated for each simulation in both directions between two of the signals, adjusting for the third. This results

in 1,000 estimated adjusted GC coefficients. The histogram of these estimates is given in Figure 2(a). The largest GC coefficient estimated is 0.0010 and the 95^{th} percentile fell at 0.0003. This process provided similar results for bivariate GC coefficients estimated from Equation (3).

The threshold choice for the real data is chosen by considering two non-overlapping time series intervals. We calculated the bivariate GC of all time series in the first interval on all time series in the second interval and vice versa for a total of $2 * 71^2$ GC estimates. The histogram of these estimates is given in Figure 2(b) . The largest GC estimated is $2.612 * 10^{-4}$ and the 95^{th} percentile fell at $4.612 * 10^{-5}$. We observe similar results when the process is repeated for adjusted GC estimates with randomly chosen time series as the adjustor.



(a) Simulated data (b) Real data

Figure 1: Null distribution of GC estimates

Based on these explorations (and the findings given in Section 5.1), all binary GC estimates exceeding 0.01 are included when identifying possible adjustors (Step 3 in Section 4.1) and 0.01 is chosen as a threshold to identify connections to be included in the adjusted GC graph (Step 6). The thresholds are not required to be equal. Choosing differing thresholds for the two steps may be justified in some situations. For example, a lower threshold could be used in Step 3 to determine which adjustors should be considered. This would result in more possible adjustors being included in the model.

5 Simulations

To examine the effectiveness of our method via simulations, we simulated 1000 white-noise, zero-mean multivariate autoregressive (MVAR) time series of order 1 and length 10,000 (after a burn-in of 1,000) using the 'arsim' function in Neumaier and Schneider's arfit toolbox for Matlab (Neumaier and Schneider, 2001; Schneider and Neumaier, 2001).

The simulations cover a variety of temporal relationships of interest. We begin with the simplest mediated relationship and feedback relationship, both with three time series. We then extend the feedback simulation to more complicated models with five time series. Based on the results of the multiplicity testing discussed in Section 4.1.2, all connections exceeding 0.01 are considered when identifying possible adjustor signals and 0.01 is also chosen as the threshold for identifying adjusted Granger causality connections. In the first simulation, we also consider the performance of the first-order pruning approach under other thresholds and incorrect model orders. We fit the models using the 'arfit' function in the arfit Matlab toolbox (Schneider and Neumaier, 2001). The models are assumed to have mean $\mathbf{0}$.

5.1 Indirect, entirely mediated

Time series described by the model in Figure 5.1 are simulated via the model given in Equation (7). Adjusted GC coefficients are estimated under several assumptions. Three combinations of thresholds and three possible model orders from Steps 3 and 6 in Section 4.1 are considered, for a total of nine settings. The fitting procedure is outlined in the first paragraph of Section 5.

$$\begin{aligned} A_t &= 0.9A_{t-1} + \epsilon_t \\ B_t &= -0.5B_{t-1} + -0.5A_{t-1} + \gamma_t \\ C_t &= 0.8C_{t-1} + 0.4B_{t-1} + \eta_t \end{aligned} \qquad \Sigma_t = \begin{bmatrix} 3.0 & 0 & 0 \\ 0 & 3.0 & 0 \\ 0 & 0 & 3.0 \end{bmatrix} \tag{7}$$

While all cutoff choices are relatively robust to misspecified model order, the conservative 0.01, 0.01 is most robust to misspecification of the order, as the correct graph is estimated

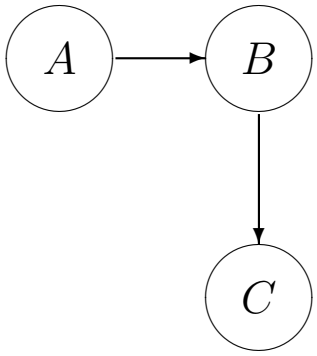
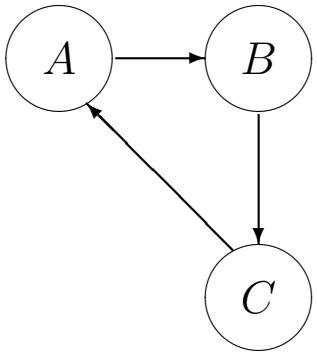


Figure 2: Graph of the model described in Equation (7)

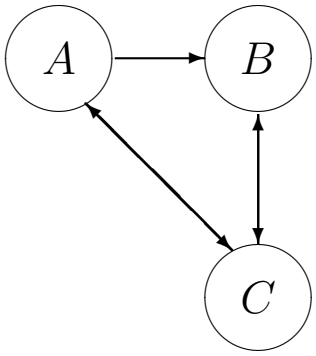
Adj, bin cutoffs	0.01, 0.01			0.001, 0.01			0.001, 0.001		
Order	**1**	5	10	**1**	5	10	**1**	5	10
A → B|C	1000	1000	1000	1000	1000	1000	1000	1000	1000
A → C|B	0	0	0	0	0	0	0	5	25
B → A|C	0	0	0	0	0	0	2	9	13
B → C|A	1000	1000	1000	1000	1000	1000	1000	1000	1000
C → A|B	0	0	0	3	10	27	0	6	27
C → B|A	0	0	0	1	10	34	2	9	17

Table 1: Each cell contains the number of simulations (out of 1000) that identified a particular adjusted connection (listed along the left) as exceeding the threshold under various model order assumptions (adjusted and binary cutoffs and order, listed along the top). The true connections and model order are in blue.

under all three orders (see Table 1). Furthermore, note that a correct connection is never left out of the graph; the only mistakes are the inclusion of spurious connections. Had we considered only bivariate GC connections in this setting, a spurious connection from A to C is estimated in all 1000 simulations and a spurious connection from B to A is estimated in 742 of the 1000 simulations in addition to the two true connections. From this point onward, we will focus on simulations where the model is fit assuming an order of 1 and the threshold is 0.01 in both Step 3 and Step 6 of the algorithm given in Section 4.1.



(a) Graph of the true model described by Equation (8) and estimated by the adjusted method in all 1,000 simulations



(b) Nearly saturated graph estimated by the bivariate method in all 1,000 simulations

Figure 3: Graphs estimated in the feedback simulations

5.2 Feedback

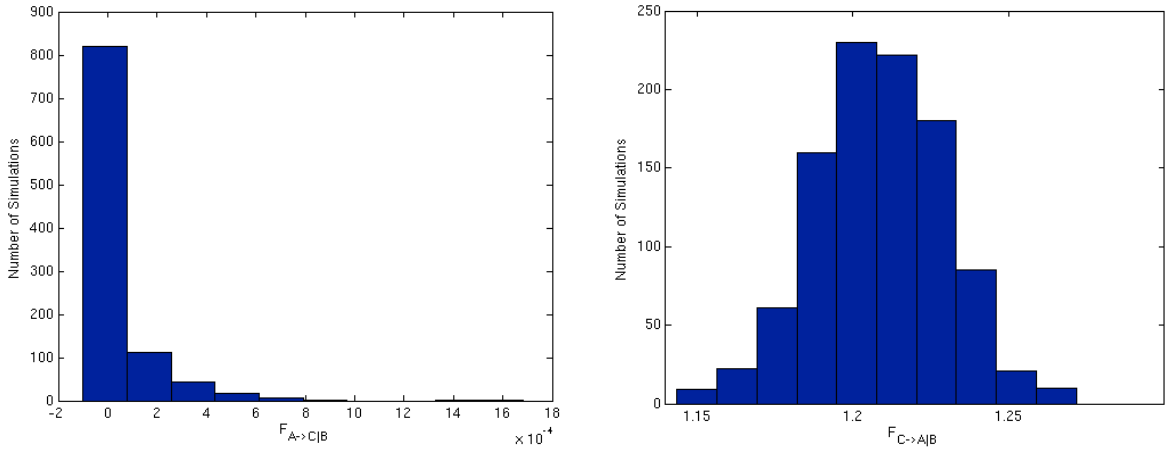
Time series described by the model in Figure 3(a) are simulated via the model given in Equation (8). Both bivariate and adjusted GC coefficients are estimated. The fitting procedure is outlined in the first paragraph of Section 5.

$$\begin{aligned} A_t &= 0.2A_{t-1} + -0.9C_{t-1} + \epsilon_t \\ B_t &= -0.8B_{t-1} + -0.9A_{t-1} + \gamma_t \\ C_t &= 0.2C_{t-1} + 0.7B_{t-1} + \eta_t \end{aligned} \qquad \Sigma_t = \begin{bmatrix} 3.0 & 0 & 0 \\ 0 & 3.0 & 0 \\ 0 & 0 & 3.0 \end{bmatrix} \tag{8}$$

The nearly saturated graph (Figure 3(b)) is estimated in each of the 1,000 simulated MVAR time series when considering only bivariate relationships. The graph estimated from the fully adjusted GC, on the other hand, is the graph given in Figure 3(a). When adjusting for the third time series, the GC method correctly identifies the direction of the feedback cycle present in the simulated data. The true model is estimated by the adjusted GC in each of the 1,000 simulations.

Each approach is consistent across the 1,000 simulations because the spread of the GC

estimates is small compared to the cutoff. Recall that a cutoff of 0.01 is applied to the simulated data to decide the presence of an edge. The histogram of the 1,000 adjusted GC estimates of $F_{A \to C|B}$ is given in Figure 4(a). No connection exists between A and C in the true model, so we would like all estimates to fall below the 0.01 threshold. In fact, all of the estimates do fall far below the 0.01 cutoff with no estimate exceeding $1.7 * 10^{-3}$. The histogram of the 1,000 adjusted GC estimates of $F_{C \to A|B}$, a connection that does exist in the true model, is given in Figure 4(b). Again, the distribution of estimates is distant from the 0.01 cutoff.



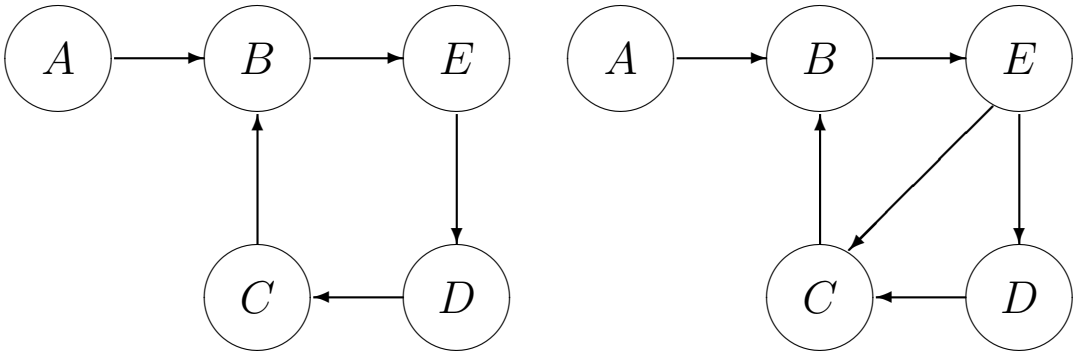
(a) $F_{A \to C|B}$ (b) $F_{C \to A|B}$

Figure 4: Histograms of GC estimates from simulations

5.3 Larger feedback model

The pruning approach to estimate directed graphs using adjusted GC successfully estimates even larger feedback graphs, such as the graph generated by the model given in Equation (9). The correct graph, given in Figure 5(a) is estimated in all but 1 of the 1,000 simulations.

However, when an edge from $E \to C$ is added to the graph (as shown in Figure 5(b) and described by the model in Equation (10)), the simulations estimate a spurious edge from E to A in approximately 1/3 of the simulations. If we were to fit the full adjusted model, i.e. a model that includes all time series as adjustors, instead of allowing only one adjustor



(a) Feedback graph defined by Equation (9) (b) Feedback graph defined by Equation (10)

Figure 5: Larger feedback graphs

per GC coefficient, then the true model is correctly estimated in all 1,000 simulations. For relatively small data sets (less than about 10 time series), we can fit the full adjusted model. For larger models, the coefficients in the autoregressive models are estimated with too much uncertainty to reliably estimate the GC coefficients (Lütkepohl, 2005).

$$\begin{aligned} A_t &= 0.9A_{t-1} + \epsilon_t \\ B_t &= -0.8B_{t-1} + 0.6A_{t-1} + 0.5C_{t-1} + \gamma_t \\ C_t &= 0.2C_{t-1} + -0.4D_{t-1} + \eta_t \\ D_t &= 0.7D_{t-1} + -0.4E_{t-1} + \xi_t \\ E_t &= 0.8E_{t-1} + -0.2B_{t-1} + \zeta_t \end{aligned} \qquad \Sigma_t = \begin{bmatrix} 3.0 & 0 & 0 & 0 & 0 \\ 0 & 3.0 & 0 & 0 & 0 \\ 0 & 0 & 3.0 & 0 & 0 \\ 0 & 0 & 0 & 3.0 & 0 \\ 0 & 0 & 0 & 0 & 3.0 \end{bmatrix} \tag{9}$$

$$
\begin{aligned}
A_t &= 0.9A_{t-1} + \epsilon_t \\
B_t &= -0.8B_{t-1} + 0.6A_{t-1} + 0.5C_{t-1} + \gamma_t \\
C_t &= 0.2C_{t-1} + -0.4D_{t-1} + 0.7E_{t-1} + \eta_t \\
D_t &= 0.7D_{t-1} + -0.4E_{t-1} + \xi_t \\
E_t &= 0.8E_{t-1} + -0.2B_{t-1} + \zeta_t
\end{aligned}
\qquad
\Sigma_t = \begin{bmatrix} 3.0 & 0 & 0 & 0 & 0 \\ 0 & 3.0 & 0 & 0 & 0 \\ 0 & 0 & 3.0 & 0 & 0 \\ 0 & 0 & 0 & 3.0 & 0 \\ 0 & 0 & 0 & 0 & 3.0 \end{bmatrix} \tag{10}
$$

6 ECoG data application

6.1 ECoG recording

Intracranial recordings were obtained from an adult, male patient with a history of medically intractable seizures who had subdural electrodes implanted over the lateral surface of the right hemisphere for clinical purposes of seizure localization and functional mapping. Event-related responses were elicited using an established 300-trial passive odd-ball paradigm with tone stimuli. 1000 Hz and 2000 Hz steady-state tones were presented sequentially and binaurally through insert earphones (70 dB SPL) at a fixed inter-stimulus interval of 1000 ms for a total of 410 seconds (Boatman and Miglioretti, 2005; Sinai et al., 2005). The patient was awake and watched a silent animated video. Event-related responses were recorded simultaneously from a total of 71 electrodes. The continuous EEG was amplified (5 X 1000) and recorded digitally (Stellate Systems, Inc.) using a referential montage, 1000-Hz A/D sampling, and a bandwidth of 0.1-300 Hz, as described previously (Sinai et al., 2009; Boatman-Reich et al., 2010). Markers for stimulus onset times were recorded simultaneously in separate marker channels. Informed written consent was obtained for the auditory recording studies in compliance with our institutional review board.

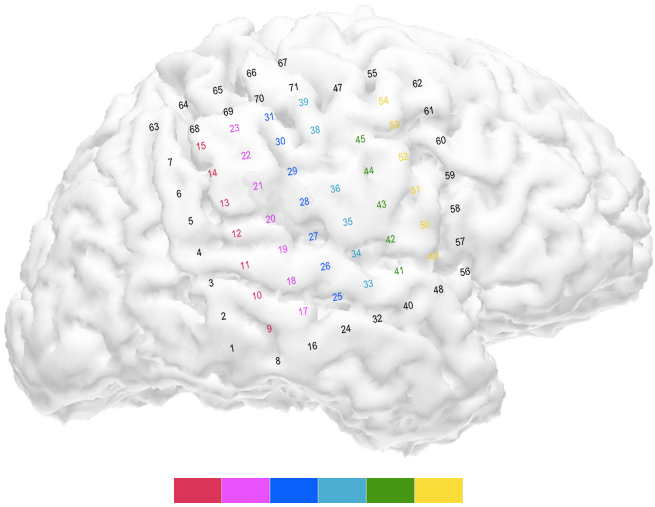


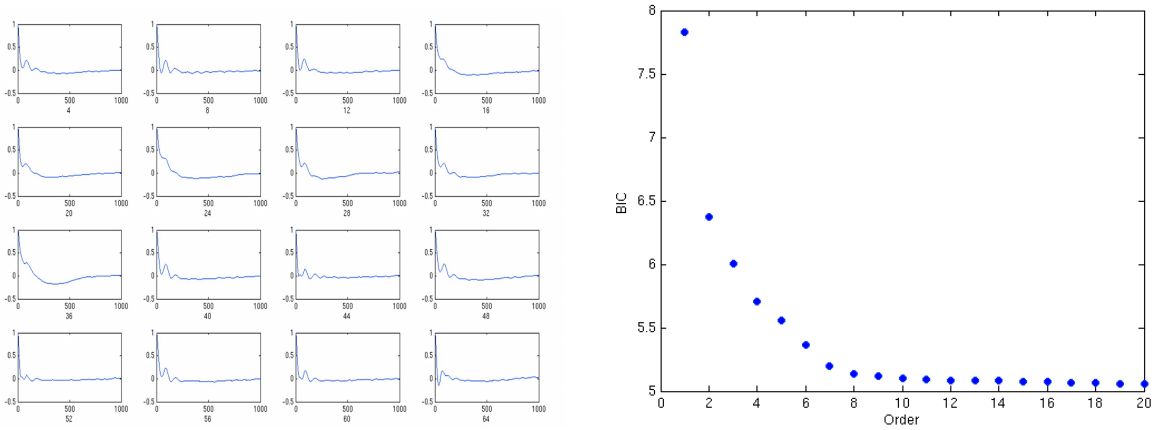
Figure 6: Co-registered 3D MRI and CT images show locations of subdural electrodes on the lateral surface of right hemisphere. Electrodes color-coded from posterior to anterior to match the node colors in Figures 8 and 9.

6.2 ECoG data analysis

The continuous EEG recording was inspected visually to identify electrodes (channels) with excessive artifact or epileptiform activity that were then excluded from analysis. Two recording channels were excluded due to artifact.

To estimate the pruned adjusted GC graph for the ECoG data, we followed the algorithm steps given in Section 4.1. A threshold of 0.01 was chosen a priori to be used in Steps 3 and 6, based on the findings given in Section 4.1.2. The AR models were estimated assuming order 7, as this model order had the lowest Schwarz's BIC for the time series pairs examined (see Figure 8(b)) (Lütkepohl, 2005). Each time series was de-meaned prior to fitting the models to reduce the number of parameters to be estimated and to improve the estimation of the covariances. The 'arfit' function included in Neumaier and Schneider's Matlab toolbox was again used to fit the AR models.

The resulting 69 x 69 matrix of 0's and 1's indicates the presence of directed, adjusted GC connections estimated by the algorithm given in Section 4.1. A by-product of the algorithm is an estimate of the directed, binary GC connections. By examining this matrix,



(a) Autocorrelations of selected electrodes

(b) Plot of Bayesian information criterion (BIC) by AR model order for the ECoG data

Figure 7: Real data diagnostics completed prior to fitting MVAR model

we found that neighboring electrodes are often estimated to have GC coefficients exceeding the threshold. This is expected due to the nature of ECoG recordings where it is common for neighboring electrodes recordings' to partially overlap. These short-distance connections are not of scientific interest in this application so we removed connections between adjacent electrodes. Furthermore, we removed the outermost electrodes and their connections from the graph (but not from the estimation of the models, i.e. we adjust for these electrodes), because we do not have recordings from their neighboring areas and thus we cannot adjust for any mediating or confounding originating outside the recording region.

The binary graph is plotted in Figure 8 and the adjusted graph is plotted in Figure 9. The plots were created using the R package Rgraphviz (Gentry et al., 2010; R Development Core Team, 2010). The nodes are color-coded in each by the location of the node's corresponding electrode on the cortex. The edges in the adjusted graph are color-coded by the color of the originating node. Figure 6 is the key to the colors and locations of the electrodes on the cortex. Black electrodes in Figure 6 were included when calculating the GC estimates but excluded when creating the graph.

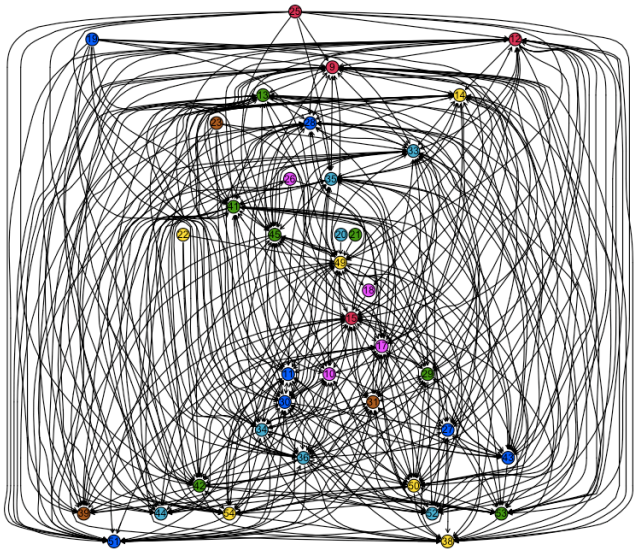


Figure 8: Graph illustrating connections estimated by bivariate GC with nodes color-coded by the corresponding electrode's brain location (refer to Figure 6)

7 Network approach to adjusted GC

Another approach to estimating adjusted GC for large sets of time series requires estimating the underlying source signals via independent components analysis (ICA). ICA is widely applied to uncover underlying source signals in EEG and ECoG data and as a dimension reduction technique Hyvärinen (1999); Hyvärinen and Oja (2000). By applying ICA to the data, we simultaneously reduce the dimensionality and identify the underlying networks, represented by the independent components. By reducing the dimensionality of the data, the fully adjusted GC model including all components as adjustors can be fit to the independent components. The resulting graph will contain information on temporal relationships between the source signals generating the raw ECoG or EEG data. Each component is a mixture of the time series observed in the original data. Thus, the information contained in the relationships between the independent components can be applied to the original signal by determining which brain locations load most strongly onto the components with connections in the resulting graph.

The ICA approach to adjusted GC is appealing because it estimates temporal associations

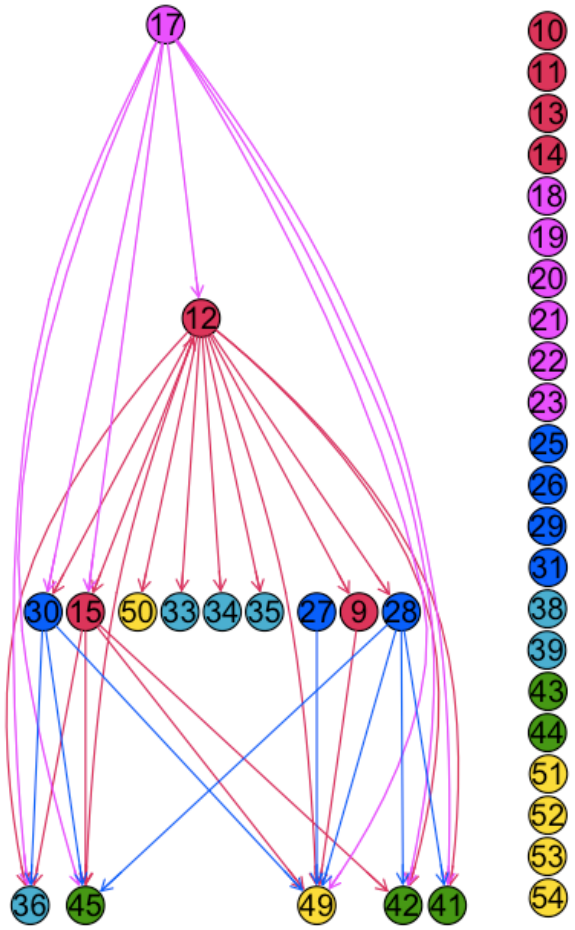


Figure 9: Graph illustrating connections estimated by adjusted GC with nodes color-coded by the corresponding electrode's brain location (refer to Figure 6) and edges color-coded according to originating node (Note that the nodes vertically aligned on the right-most portion of the graph are singletons, i.e. nodes that have no connections)

between the underlying source signals, which is the primary scientific question of interest. Furthermore, the dimension reduction in ICA permits the ideal, fully adjusted GC model to be fit to the independent components. Simulations to support the ICA approach are forthcoming.

8 Discussion

We propose a computationally feasible approach to adjust for mediating and confounding when estimating temporal associations between time series. In lieu of fitting the ideal, yet inestimable, fully adjusted model we fit a first-order pruned model to an sample ECoG dataset with 69 time series using the algorithm provided in Section 4.1. Even when considering a single adjustor between each pair of time series, we see striking improvements in the first-order adjusted GC graph (Figure 9) over the unadjusted bivariate GC graph (Figure 8). Any important connections in the bivariate GC graph are being overwhelmed by spurious connections. Patterns that may exist within the graph are hidden by the extensive number of connections. By allowing for adjustment, the number of connections is greatly reduced, leaving only those connections that are direct and not solely due to confounding or mediating signals. Furthermore, the pruning approach to GC is not restricted in the number of time series that may be simultaneously considered in the estimation because each fitted MVAR is only of dimension three in the first-order adjusted GC model.

We adjust for possible confounding and mediating associations between time series to avoid identifying spurious connections. However, our approach considers only observed time series; hence, it is possible that adjusted connections estimated using this method may be mediated or confounded by unobserved latent time series. Connections between time series that lie near the borders of the recorded region are more susceptible to being temporally associated with unobserved signals while connections between time series that lie in the interior of the recorded region are less likely to be affected by external signals. In the real data application, we removed connections involving electrodes lying on the boundary of the recording region to avoid spurious connections due to unobserved signals.

Currently, most Granger causal analyses are restricted to linear relationships between time series and this work follows in this practice. Recent works have considered the possibility of estimating nonlinear relationships using GC approaches (Marinazzo et al., 2010; Dhamala et al., 2008; Chávez et al., 2003, for example). Others have considered GC methods that

permit time-varying coefficients (Sato et al., 2010; Hesse et al., 2003). The ECoG data example presented above does not consider nonlinear, time variant, latent, or exogenous signals; however, the pruning approach we propose can be applied to any extension of GC that permits applying a threshold to fitted bivariate and adjusted GC model estimates.

In the future we hope to extend the pruning approach to allow statistical comparisons across subjects and/or repeated measures. This advance will present an opportunity to formalize statistical comparisons across subjects such as tests of differences between subjects or over time, identification of a discrimination tool between normal and impaired patients, etc. The statistical comparisons could be extensions of traditional statistics or more novel statistics such as graph theory measures like clustering coefficients or average path length (Brandes and Erlebach, 2005). The ICA concept mentioned above also holds promise as a technique for across subject comparisons. The pruning method presented within for estimating Granger causality between many time series compromises between computational feasibility and adequately adjusting for observed signals, while remaining sufficiently general to incorporate extensions proposed by other authors.

References

Arnold, M., Milner, X., Witte, H., Bauer, R., and Braun, C. (1998). Adaptive AR modeling of nonstationary time series by means of Kalman filtering. *Biomedical Engineering, IEEE Transactions on*, 45(5):553–562.

Astolfi, L., Cincotti, F., Mattia, D., Marciani, M. G., Baccala, L. A., de Vico Fallani, F., Salinari, S., Ursino, M., Zavaglia, M., Ding, L., Edgar, J. C., Miller, G. A., He, B., and Babiloni, F. (2007). Comparison of different cortical connectivity estimators for high-resolution EEG recordings. *Human Brain Mapping*, 28(2):143–157.

Blinowska, K. J., Kuś, R., and Kamiński, M. (2004). Granger causality and information flow in multivariate processes. *Phys. Rev. E*, 70(5):050902.

Boatman, D. F. and Miglioretti, D. L. (2005). Cortical sites critical for speech discrimination in normal and impaired listeners. *J. Neurosci.*, 25(23):5475–5480.

Boatman-Reich, D., Franaszczuk, P., Korzeniewska, A., Caffo, B., Ritzl, E., Colwell, S., and Crone, N. (2010). Quantifying auditory event-related responses in multichannel human intracranial recordings. *Frontiers in Computational Neuroscience*, 4:4.

Brandes, U. and Erlebach, T. (2005). *Network Analysis: methodological foundations*. Springer Verlag.

Bressler, S. and Ding, M. (2002). Event-related potentials. In *The handbook of brain theory and neural networks*, pages 412–415. John Wiley & Sons, Inc.

Bressler, S. L. and Seth, A. K. (2010). Wiener-granger causality: A well established methodology. *NeuroImage*, In Press, Corrected Proof.

Bullmore, E. and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3):186–198.

Chávez, M., Martinerie, J., and Quyen, M. L. V. (2003). Statistical assessment of nonlinear causality: application to epileptic EEG signals. *Journal of Neuroscience Methods*, 124(2):113 – 128.

Chen, Y., Rangarajan, G., Feng, J., and Ding, M. (2004). Analyzing multiple nonlinear time series with extended Granger causality. *Physics Letters A*, 324(1):26–35.

Dauwels, J., Vialatte, F., Musha, T., and Cichocki, A. (2009). A comparative study of synchrony measures for the early diagnosis of Alzheimer's disease based on EEG. *Neuroimage*, 49(1):668–693.

David, O., Harrison, L., and Friston, K. J. (2005). Modelling event-related responses in the brain. *NeuroImage*, 25(3):756 – 770.

David, O., Kiebel, S. J., Harrison, L. M., Mattout, J., Kilner, J. M., and Friston, K. J. (2006). Dynamic causal modeling of evoked responses in EEG and MEG. *NeuroImage*, 30(4):1255 – 1272.

Dhamala, M., Rangarajan, G., and Ding, M. (2008). Analyzing information flow in brain networks with nonparametric Granger causality. *NeuroImage*, 41(2):354 – 362.

Ding, M., Bressler, S., Yang, W., and Liang, H. (2000). Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: data preprocessing, model validation, and variability assessment. *Biological cybernetics*, 83(1):35–45.

Ding, M., Chen, Y., and Bressler, S. L. (2006). Granger causality: Basic theory and application to neuroscience. In Schelter, B., Winterhalder, M., and Timmer, J., editors, *Handbook of time series analysis: recent theoretical developments and applications*, pages 451–474. Wiley-VCH Verlage.

Eichler, M. (2005). A graphical approach for evaluating effective connectivity in neural systems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1457):953–967.

Friston, K. (1994). Functional and effective connectivity in neuroimaging: a synthesis. *Human Brain Mapping*, 2.

Friston, K. (2004). *Models of brain function in neuroimaging*. Annual Reviews.

Gao, W. and Tian, Z. (2009). Learning Granger causality graphs for multivariate nonlinear time series. *Journal of Systems Science and Systems Engineering*, 18(1):38–52.

Gentry, J., Long, L., Gentleman, R., Falcon, S., Hahne, F., Sarkar, D., and Hansen, K. (2010). *Rgraphviz: Plotting capabilities for R graph objects*. R package version 1.26.0.

Geweke, J. (1982). Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association*, 77(378):304–313.

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

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438.

Guo, S., Seth, A. K., Kendrick, K. M., Zhou, C., and Feng, J. (2008). Partial Granger causality–eliminating exogenous inputs and latent variables. *Journal of Neuroscience Methods*, 172(1):79–93.

Havlicek, M., Jan, J., Brazdil, M., and Calhoun, V. D. (2010). Dynamic granger causality based on kalman filter for evaluation of functional network connectivity in fmri data. *NeuroImage*, In Press, Corrected Proof.

Hesse, W., Moller, E., Arnold, M., and Schack, B. (2003). The use of time-variant EEG Granger causality for inspecting directed interdependencies of neural assemblies. *Journal of Neuroscience Methods*, 124(1):27 – 44.

Horwitz, B. (2003). The elusive concept of brain connectivity. *NeuroImage*, 19(2):466 – 470.

Hyvärinen, A. (1999). Survey on independent component analysis. *Neural Computing Surveys*, 2(4):94–128.

Hyvärinen, A. and Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural networks*, 13(4-5):411–430.

Jiao, Q., Lu, G., Zhang, Z., Zhong, Y., Wang, Z., Guo, Y., Li, K., Ding, M., and Liu, Y. (2010). Granger causal influence predicts BOLD activity levels in the default mode network. *Human Brain Mapping*.

Kamiński, M., Ding, M., Truccolo, W., and Bressler, S. (2001). Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. *Biological Cybernetics*, 85(2):145–157.

Lantz, G., de Peralta, R. G., Spinelli, L., Seeck, M., and Michel, C. M. (2003). Epileptic source localization with high density EEG: how many electrodes are needed? *Clinical Neurophysiology*, 114(1):63 – 69.

Lauritzen, S. (1996). *Graphical models*. Oxford University Press, USA.

Liao, W., Mantini, D., Zhang, Z., Pan, Z., Ding, J., Gong, Q., Yang, Y., and Chen, H. (2010). Evaluating the effective connectivity of resting state networks using conditional Granger causality. *Biological Cybernetics*, 102(1):57–69.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer.

Marinazzo, D., Liao, W., Chen, H., and Stramaglia, S. (2010). Nonlinear connectivity by Granger causality. *NeuroImage*, In Press, Corrected Proof.

Neumaier, A. and Schneider, T. (2001). Estimation of parameters and eigenmodes of multivariate autoregressive models. *ACM Transactions on Mathematical Software (TOMS)*, 27(1):27–57.

Oostenveld, R. and Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology*, 112(4):713–719.

Oya, H., Poon, P. W., Brugge, J. F., Reale, R. A., Kawasaki, H., Volkov, I. O., and III, M. A. H. (2007). Functional connections between auditory cortical fields in humans revealed by Granger causality analysis of intra-cranial evoked potentials to sounds: Comparison of two methods. *Biosystems*, 89(1-3):198 – 207.

R Development Core Team (2010). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Sato, J. R., Fujita, A., Cardoso, E. F., Thomaz, C. E., Brammer, M. J., and Jr., E. A. (2010). Analyzing the connectivity between regions of interest: An approach based on cluster Granger causality for fMRI data analysis. *NeuroImage*, 52(4):1444 – 1455.

Schneider, T. and Neumaier, A. (2001). Algorithm 808: ARfit - A Matlab package for the estimation of parameters and eigenmodes of multivariate autoregressive models. *ACM Transactions on Mathematical Software (TOMS)*, 27(1):58–65.

Shaw, J. (1981). An introduction to the coherence function and its use in EEG signal analysis. *Journal of Medical Engineering & Technology*, 5(6):279–288.

Sinai, A., Bowers, C., Crainiceanu, C., Boatman, D., Gordon, B., Lesser, R., Lenz, F., and Crone, N. (2005). Electrocorticographic high gamma activity versus electrical cortical stimulation mapping of naming. *Brain*, 128(7):1556–1570.

Sinai, A., Crone, N., Wied, H., Franaszczuk, P., Miglioretti, D., and Boatman-Reich, D. (2009). Intracranial mapping of auditory perception: Event-related responses and electrocortical stimulation. *Clinical Neurophysiology*, 120(1):140 – 149.

Walter, D. O. (1963). Spectral analysis for electroencephalograms: Mathematical determination of neurophysiological relationships from records of limited duration. *Experimental Neurology*, 8(2):155 – 181.

Wiener, N. (1956). The theory of prediction. In Beckenback, E. F., editor, *Modern mathematics for engineers*, 1, pages 125–139.