

HHS Public Access

Author manuscript *Neural Comput.* Author manuscript; available in PMC 2017 August 26.

Published in final edited form as:

Neural Comput. 2016 May ; 28(5): 914–949. doi:10.1162/NECO_a_00828.

State-Space Analysis of Granger-Geweke Causality Measures with Application to fMRI

Victor Solo

School of Electrical Engineering, UNSW, Sydney, Australia, and MGH/HST Martinos Center for Biomedical Imaging, Department of Radiology-MGH, Harvard Medical School, Boston, MA, U.S.A

Abstract

The recent interest in the dynamics of networks and the advent, across a range of applications, of measuring modalities that operate on different temporal scales have put the spotlight on some significant gaps in the theory of multivariate time series. Fundamental to the description of network dynamics is the direction of interaction between nodes, accompanied by a measure of the strength of such interactions. Granger causality and its associated frequency domain strength measures (GEMs) (due to Geweke) provide a framework for the formulation and analysis of these issues. In pursuing this setup, three significant unresolved issues emerge. First, computing GEMs involves computing submodels of vector time series models, for which reliable methods do not exist. Second, the impact of filtering on GEMs has never been definitively established. Third, the impact of downsampling on GEMs has never been established. In this work, using state-space methods, we resolve all these issues and illustrate the results with some simulations. Our analysis is motivated by some problems in (fMRI) brain imaging, to which we apply it, but it is of general applicability.

1 Introduction

Following the operational development of the notion of (temporal) causality by Granger (1969) and Sims (1972), Granger causality (GC) analysis has become an important part of time series and econometric testing and inference (Hamilton, 1994). It has also been applied in the biosciences (Kaminski & Blinowska, 1991; Bernasconi & Konig, 1999; Ding, Bressler, Yang, & Liang, 2000), climatology (global warming; Suna & Wang, 1996; Kaufmann & Stern, 1997; Triacca, 2005), and, most recently, functional magnetic resonance imaging (fMRI). This last application has stimulated this letter.

Since its introduction into fMRI (Goebel, 2003; Valdes-Sosa, 2004; Roebroeck, Formisano, & Goebel, 2005; Yamashita, Sadato, Okada, & Ozaki, 2005) GC has become the subject of an intense debate: (see Valdes-Sosa, Roebroeck, Daunizeau, & Friston, 2011; Roebroeck, Formisano, & Goebel, 2011) and associated commentary. There are two main issues in that debate that occur more widely in dynamic networks. First is the impact of downsampling on GC. In the fMRI neuroimaging application, causal processes may operate on a timescale on the order of tens of milliseconds, whereas the recorded signals are typically available only on a 1 second timescale, although sampling up to a 100 ms rate is now becoming possible

(Feinberg & Yacoub, 2012). So it is natural to wonder if GC analysis on a slower timescale can reveal GC structure (GCS) on a faster timescale. Second is the impact of filtering on GC due to the hemodynamic response function, which relates neural activity to the recorded fMRI signal. Since intuitively GC will be sensitive to time delay, the variability of the hemodynamic response function (HRF), particularly spatially varying delay (called time to onset in the fMRI literature) and time to peak (confusingly sometimes called delay in the fMRI literature) has been suggested as a potential source of problems (Deshpande, Sathian, & Hu, 2010; Handwerker, Gonzalez-Castillo, D'Esposito, & Bandettini, 2012). A referee has, however, pointed out that following David et al. (2008), deconvolution (Glover, 1999) has gained favor (Havlicek, Friston, Jan, Brazdil, & Calhoun, 2011) as a potential means of dealing with this.

An important advance in GC theory and tools was made by Geweke (1982), who provided measures of the strength of causality or influence of one time series on another (henceforth called GEM, for Geweke causality measure), including frequency domain decompositions of them. Subsequently it was pointed out that the GEMs are measures of directional mutual information (Rissannen & Wax, 1987). The GEMs were extended to conditional causality in Geweke (1984). However, GEMs have not found as wide an application as they should have, partly because of some technical difficulties in calculating them, which are discussed further below. But GEMs (and their frequency domain versions) are precisely the tool needed to pursue both GC downsampling and filtering questions.

In the econometric literature, it was appreciated early that downsampling, especially in the presence of aggregation, could cause problems. This was implicit in the work of Sims (1971) and mentioned also in the work of Christiano and Eichenbaum (1987), who gave an example of contradictory causal analysis based on monthly versus quarterly data, and also discussed in Marcet (1991). But precise general conditions under which problems do and do not arise have never been given. We do so in this letter.

Some of the econometric discussion has been framed in terms of sampling of continuous time models (Sims, 1971; Marcet, 1991; Christiano & Eichenbaum, 1987). And authors such as Sims (1971) have suggested that models are best formulated initially in continuous time. While this is a view that I have long shared (and related to the modeling approach advocated by Middleton and Goodwin (1990), we deal with only discrete time models here. To cast our development in terms of continuous time models would require a considerable development of its own without changing our basic message.

The issue at stake, in its simplest form, is the following. Suppose that a pair of (possibly vector) processes exhibit a unidirectional GC relation (a influences b, but b does not influence a), but suppose measurement time series are available only at a slower timescale or as filtered time series or both. Then two questions arise. The first, which we call the *forward* question, is this: When is the unidirectional GC relation preserved? The second, which we call the *reverse* question, is harder. Suppose the measurement time series exhibit a unidirectional GC relation. Does that mean the underlying unfiltered faster timescale processes do? The latter question is the more important and so far has received no theoretical

attention. The more general form of this issue is to what extent the GC structure (GCS) between two time series is distorted by downsampling and or filtering or both.

In order to resolve these issues, we need to develop some theory and some computational and modeling tools. First, to compute GEMs, one needs to be able to find submodels from a larger model (i.e., one having more time series). Thus, to compute the GEMs between time series $x_b y_b$ Geweke (1982, 1984) attempted to avoid this by fitting submodels separately to x_t to y_t and then also fitting a joint model to $x_b y_t$. Unfortunately this can generate negative values for some of the frequency domain GEMs (Chen, Bressler, & Ding, 2006). Properly computing submodels will resolve this problem, and previous work has not accomplished this (we discuss the attempts in Dufour & Taamouti, 2010, and Chen et al., 2006, below).

Second, one needs to be able to compute how models transform when they are downsampled. This has been done only in special cases (Pandit & Wu, 1983) or by methods that are not computationally realistic. We provide computationally reliable, state-space-based methods for doing this here.

Third, we need to study the effect of downsampling and filtering on GEMs. We will show how to use the newly developed tools to do this.

To sum up, we can say that previous discussions, including those above, as well as Geweke (1978), Telser (1967), Sims (1971), and Marcet (1991), fail to provide general algorithms for finding submodels or models induced by downsampling. Indeed both these problems have remained open problems in multivariate time series in their own right for several decades, and we resolve them here. Further, there does not seem to have been any theoretical discussion of the effect of filtering on GEMs, and we resolve that here also. To do that, it turns out that state-space models provide the proper framework.

Throughout this work, we deal with the dynamic interaction between two vector time series. It is well known in econometrics that if there is a third vector time series involved in the dynamics but not accounted for, then spurious causality can occur for reasons that have nothing to do with downsampling, a situation discussed by Hsiao (1982) and Geweke (1984). Other causes of spurious causality such as observation noise are also not discussed. Of course, the impact of downsampling in the presence of a third (vector) variable is also of interest but will be pursued elsewhere.

Finally, our whole discussion is carried out in the framework of linear time series models. It is of great interest to pursue nonlinear versions of these issues, but that would be a major separate task.

The remainder of the letter, is organized as follows. In section 2, we review and modify some state-space results important for system identification or model fitting and needed in the following sections. In section 3, we develop state-space methods for computing submodels of innovations state-space models. In section 4, we develop methods for transforming state-space models under downsampling. In section 5, we review GC and GEMs and extend them to a state space-setting. In section 6, we study the effect of filtering on GC via frequency domain GEMs. In section 7, we offer a theory to explain when GC is

preserved under downsampling. In section 8, we discuss the reverse GC problem showing how spurious causality can be induced by downsampling. In section 9, we set up the framework for the application of these results to fMRI. This includes discussion of the phase properties of HRFs as well as the impact of deconvolution on the assessment of GC. Conclusions are in section 10. There are three appendixes.

1.1 Acronyms and Notation

GC is Granger causality or Granger causes; *GCS* is Granger causal structure; *dn-gc* does not Granger cause; *GEM* is Geweke causality measure. *HRF* is hemodynamic response function; *SS* is state space or state-space model; *ISS* is innovations state-space model; *UC* is unit circle; *VAR* is vector autoregression; *VARMA* is vector autoregressive moving average process; *wp1* is with probability 1. *LHS* denotes left-hand side. The *lag* or backshift operator is denoted *L* or z^{-1} ; thus, if x_t is a signal, then $Lx_t = z^{-1} x_t = x_{t-1}$. *L* or z^{-1} are causal operators since they operate on the past. L^{-1} or *z* are noncausal operators since they operate

on the future (so $L^{-1}x_t = zx_t = x_{t+1}$). The causal filter $h(L) = \sum_{0}^{\infty} h_r L^r$ is sometimes denoted *h*, and its noncausal reflection is $h^* = h^*(L) = h(L^{-1})$.

 X_a^b denotes the values x_a , x_{a+1} , ..., x_b , so $X_a^a \equiv x_a$. For stationary processes, we have $a = -\infty$. If *M*, *N* are positive semidefinite matrices, then *M N* means *M*–*N* is positive semidefinite. A *stable* square matrix is one whose eigenvalues all have modulus < 1.

2 State Space

The computational methods we develop rely on state-space techniques and spectral factorization, the latter being intimately related to the steady-state Kalman filter. In this section, we review and modify some basic results in state-space theory, Kalman filtering, and spectral factorization. Our discussion deals with two vector time series, collected as,

$$z_t = \left(x_t^T, y_t^T\right)^T.$$

2.1 State-Space Models

We consider a general constant parameter SS model,

$$\xi_{t+1} = A\xi_t + w_t, \ z_t = C\xi_t + v_t, \ (2.1)$$

with positive semidefinite noise covariance, $var\begin{pmatrix} w_t \\ v_t \end{pmatrix} = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix}$. We refer to this as an SS model with parameters (A, C, [Q, R, S]).

It is common with SS models to take S = 0, but for equivalence between the class of VARMA models and the class of state-space models, it is necessary to allow S = 0.

Now by matrix partitioning, $\begin{vmatrix} Q & S \\ S^T & R \end{vmatrix} = |R| |Q_s|, Q_s = Q - SR^{-1}S^T$. So we introduce

N: Noise Condition. *R* is positive definite.

N implies that Q_s is positive semidefinite.

2.2 Steady-State Kalman Filter, Innovations State-Space Models, and the Discrete Algebraic Ricatti Equation

We now recall the Kalman filter for mean square estimation of the unobserved state sequence ξ_t from the observed time series z_t . It is given by Kailath, Sayeed, and Hassibi (2000, theorem 9.2.1),

 $\hat{\xi}_{t+1} = A\hat{\xi}_t + K_t e_t, \ e_t = z_t - C\hat{\xi}_t, \ \text{or } z_t = C\hat{\xi}_t + e_t,$

where e_t is the innovations sequence of variance $V_t = R + CP_tC^T$, $K_t = (AP_tC^T + S)V_t^{-1}$ is the Kalman gain sequence, and P_t is the state error variance matrix generated from the Ricatti equation, $P_{t+1} = AP_tA^T + Q - K_tV_tK_t^T$.

The Kalman filter is a time-varying filter, but we are interested in its steady state. If there is a steady state (i.e., $P_t \rightarrow P$ as $t \rightarrow \infty$) then the limiting state error variance matrix *P* will, obey the so-called discrete algebraic Ricatti equation (DARE),

$$P = APA^T + Q - KVK^T, \quad (2.2)$$

where $V = R + CPC^T$ and $K = (APC^T + S)V^{-1}$ is the corresponding steady-state Kalman gain. With some clever algebra, (Kailath et al., 2000), the DARE can be rewritten (the Ricatti equation can be similarly rewritten) as

$$P = A_s P A_s^T + Q_s - K_s V K_s^T,$$

where $A_s = A - SR^{-1}C$ and $K_s = A_sPC^TV^{-1}$.

We now introduce two assumptions.

- St: Stabilizability condition: A_s , $O_s^{\frac{1}{2}}$ is stabilizable (see appendix A)
- De: Detectability condition: A_s , C is detectable.

In appendix A, it is shown De is equivalent to A, C being detectable. Also it holds automatically if A is stable.

The resulting steady-state Kalman filter can be written as

$$\hat{\xi}_{t+1} = A\hat{\xi}_t + K\varepsilon_t, \quad z_t = C\hat{\xi}_t + \varepsilon_t, \quad (2.3)$$

where ε_t is the steady-state innovation process and has variance matrix *V* and Kalman gain *K*. This steady-state filter provides a new state-space representation of the data sequence. We refer to it as an innovations state-space (ISS) model with parameters (*A*, *C*, *K*, *V*). We summarize this:

Result 1—Given the SS model, equation 2.1, with parameters (*A*, *C*, [*Q*, *R*, *S*]), then provided N, St, De hold:

- **a.** The corresponding ISS model, equation 2.3, with parameters (*A*, *C*, *K*, *V*), can be found by solving the DARE, equation 2.2, which has a unique positive-definite solution *P*.
- **b.** *V* is positive definite, (A, C) is detectable, and A-KC is stable so that (A, K) is controllable.

Proof—See appendix A.

Remarks

- i. Henceforth an ISS model with parameters (*A*, *C*, *K*, *V*) will be required to have *V* positive definite, (*A*, *C*) detectable, and (*A*, *K*) controllable so that *A*–*KC* is stable.
- ii. It is well known that any VARMA model can be represented as an ISS model and vice versa (Solo, 1986; Hannan & Deistler, 1988).
- iii. Note that the ISS model with parameters (*A*, *C*, *K*, *V*) can also be written as the SS model with parameters (*A*, *C*, [*KVK*^T, *V*, *KV*]).
- iv. The DARE is a quadratic matrix equation, but it can be computed using the (numerically reliable) DARE command in Matlab as follows. Compute: [P, L_0 , G] = $DARE(A^T, C^T, Q, R, S, I)$ and then, $V = R + CPC^T, K = G^T$.
- v. Note that stationarity is not required for this result.

2.3 Stationarity and Spectral Factorization

Given an ISS model with parameters (A, C, B, Σ_e), we now introduce

• Assumption EV: Eigenvalue stability condition: *A* has all eigenvalues with modulus < 1, that is, *A* is a stability matrix.

With this assumption we can obtain an infinite vector moving average (VMA) representation, an infinite vector autoregressive (VAR) representation, and a spectral factorization.

Before stating these results, we need a definition;

• Definition. Minimum phase transfer function. A matrix H(L) of causal transfer functions is said to be minimum phase (MP) if its inverse exists and is causal and stable. If H(L) is not minimum phase, it is called nonminimum phase (NMP).

Examples

i. Causal and stable. H(L) = 1 - aL, |a| < 1 has inverse

 $H_{-}(L) = \frac{1}{1-aL} = \sum_{0}^{\infty} a^{r} L^{r}$, which follows from summing the geometric series. Then $H_{-}(L)$ is clearly causal and is stable since its pole $z_{p} = L_{p}^{-1} = a$ has modulus < 1 (i.e., it lies inside the unit circle). Another way to state this is to say that H(L) is MP because its zero $z_{o} = a$ has modulus < 1 (i.e., it lies inside the unit circle, UC).

ii.

Causal but not stable. $H(L) = 1 - \frac{1}{\theta}L, 0 < \theta < 1$ is not minimum phase since $H_{-}(L) = \frac{1}{1 - \frac{1}{\theta}L}$ is not stable since its pole $z_p = L_p^{-1} = \frac{1}{\theta}$ has modulus > 1. If we rewrite $H(L) = -\frac{1}{\theta}L(1 - \theta L^{-1})$, then

 $H_{-}(L) = -\frac{\theta L^{-1}}{1 - \theta L^{-1}} = -\theta L^{-1} \sum_{0}^{\infty} \theta^{r} L^{-r}$. This is stable but not causal since the expansion is in terms of the future. Again we can rephrase this example to

say H(L) is NMP because its zero $z_o = \frac{1}{\theta}$ is outside the unit circle.

iii. Not causal but stable. H(L) = L(1 - aL), |a| < 1 is nonminimum phase since

 $H_{-}(L) = \frac{L^{-1}}{1 - aL}$ is not causal. Note that $H_{-}(L)$ is stable. We can rephrase this as follows. H(L) has two zeros: one is inside the unit circle, and the other is at ∞ (i.e., $z = L^{-1} = \infty$ makes H(L) vanish and so is outside the unit circle).

For further discussion of minimum phase filters, see Green (1988) and Solo (1986).

The following result is based on Kailath et al. (2000, theorem 8.3.2) and the surrounding discussion.

<u>Result 2</u>: For the ISS model (A, C, B, Σ_e) obeying condition Ev we have,

a. Infinite VMA or Wold decomposition:

$$z_t = H(L)\varepsilon_t = (C(L^{-1}I - A)^{-1}B + I)\varepsilon_t = (C(I - AL)^{-1}BL + I)\varepsilon_t.$$
 (2.4)

b. Infinite VAR representation:

$$\varepsilon_t = G(L)z_t = \left[I + C(L^{-1}I - A + KC)^{-1}K\right]z_t.$$
 (2.5)

c. Spectral factorization. Put $L = \exp(-j\lambda)$; then z_t has positive-definite spectrum with spectral factorization as follows:

$$f_{Z}(\lambda) = \begin{bmatrix} C(L^{-1}I - A)^{-1}, I \end{bmatrix} \begin{bmatrix} Q & S \\ S^{T} & R \end{bmatrix} \begin{bmatrix} (LI - A^{T})^{-1}C^{T} \\ I \end{bmatrix}$$
$$= H(L) \sum_{\varepsilon} H^{T}(L^{-1}).$$
(2.6)

d. H(L) is minimum phase.

Proof

- **a.** Write equation 2.3, in operator form. The series is convergent wp1 and in mean square since *A* is stable.
- **b.** Rewrite equation as $\hat{\xi}_{t+1} = (A KC) \hat{\xi}_t + Kz_t, \varepsilon_t = z_t C\hat{\xi}_t$. Then write this in operator form. The series is convergent wp1 and in mean square since A KC is stable and z_t is stationary.
- **c.** This follows from standard formulas for spectra of filtered stationary time series applied to (a).
- **d.** From (a),(b) $G(L) = H^{-1}(L)$ and by (b) G(L) is causal and stable, and the result follows.

Remarks

- i. Result 2 is a special case of a general result that given a full-rank multivariate spectrum $f_Z(\lambda)$, there exists a unique causal stable minimum phase spectral factor H(L) with H(0) = I and positive-definite innovations variance matrix Σ_e such that equation 2.6 holds (Hannan & Deistler, 1988; Green, 1988). In general *detH*(*L*) may have some roots on the unit circle (Hannan & Poskitt, 1988; Green, 1988), but the assumptions in result 2 rule this case out. Such roots mean that some linear combinations of z_t can be perfectly predicted from the past (Hannan & Poskitt, 1988; Green, 1988), something that is not realistic in the fMRI application.
- Result 2 is also crucial from a system identification or model-fitting point of view. From that point of view, all we can know (from second-order statistics) is the spectrum, and so if, naturally, we want a unique model, the only model we can obtain is the causal stable minimum phase model: the ISS model. The standard approach to SS model fitting is the so-called state-space subspace method (Deistler, Peternell, & Scherer, 1995; Bauer, 2005), and indeed it delivers an ISS model. The alternative approach of fitting a VARMA model (Hannan & Deistler, 1988; Lutkepohl, 1993) is equivalent to getting an ISS model.
- **iii.** We need result 1, however, since when we form submodels, we do not immediately get an ISS model; rather, we must compute it.

3 Submodels

Solo

Our computation of causality measures requires that we compute induced submodels. In this section, we show how to obtain an ISS submodel from the ISS joint model.

We partition $z_t = (x_b, y_t)^T$ into subvector signals of interest and partition the state-space

model correspondingly: $C = \begin{pmatrix} C_X \\ C_Y \end{pmatrix}$ and $B = (B_X, B_Y)$. We first read out an SS submodel for x_t from the ISS model for z_t . We have simply $\xi_{t+1} = A\xi_t + w_t$, $x_t = C_X\xi_t + \varepsilon_{X,t}$, where,

 $w_{t}=B\begin{pmatrix}\varepsilon_{X,t}\\\varepsilon_{Y,t}\end{pmatrix}=B_{X}\varepsilon_{X,t}+B_{Y}\varepsilon_{Y,t}.$ We need to calculate the covariance matrix: $var\begin{pmatrix}w_{t}\\\varepsilon_{X,t}\end{pmatrix}=\begin{pmatrix}Q\\\overline{S}^{T}&\overline{R}\end{pmatrix}.$ We find $\overline{R}=\sum_{X,e}, Q=var(w_{t})=B\Sigma_{e}B^{T}$ and $\overline{S}=E(w_{t}\varepsilon_{X,t}^{T})=B(\sum_{YX,e}^{\Sigma_{X,e}})=B_{o}.$ This leads to:

Theorem 1

Given the joint ISS model, equation 2.3 or 2.4, for z_b then under condition Ev, the corresponding ISS submodel for x_b (A, C_X , $K_{(X)}$, Ω_X) (the bracket notation $K_{(X)}$ is used to avoid confusion with e.g. C_X , $\Sigma_{X,e}$ which are submatrices), can be found by solving the

DARE, equation 2.2, with
$$\left[Q, \overline{R}, \overline{S}\right] = \left[B\sum_{\varepsilon} B^T, \sum_{X, \varepsilon}, B_o\right]$$
.

Proof

First, we note by partitioning $|\Sigma_{e}| = |\Sigma_{X,e}| |\Sigma_{(Y|X),e}|$ where $\sum_{(Y|X),\varepsilon} = \sum_{YX,\varepsilon} \sum_{X,\varepsilon} \sum_{X,\varepsilon} \sum_{XY,\varepsilon} \sum_{XY,$

The first is already established. The second follows trivially since A is stable. We use the PBH test (see appendix A) to check the third.

Suppose controllability fails. Then by the PBH test, there exists q = 0 with

 $\lambda q^T = q^T (A - \overline{SR}^{-1}C_x)$ and

$$\begin{split} 0 &= q^T \left(B \sum_{\varepsilon} B^T - \overline{SR}^{-1} \overline{S}^T \right)^{\frac{1}{2}} \\ \Rightarrow 0 &= q^T \left(B \sum_{\varepsilon} B^T - \overline{SR}^{-1} \overline{S}^T \right) \\ = &q^T \left(B \sum_{\varepsilon} B^T - B \left(\sum_{X,\varepsilon} \sum_{YX,\varepsilon} \right) \sum_{X,\varepsilon}^{-1} \left[\sum_{X,\varepsilon}, \sum_{XY,\varepsilon} \right] B^T \right) \\ &= &q^T \left(B_X, B_Y \right) \begin{pmatrix} 0 & 0 \\ 0 & \sum_{\left(Y \mid X\right),\varepsilon} \right) \left(\begin{array}{c} B_X^T \\ B_Y^T \end{array} \right) \\ &= &q^T B_Y \sum_{\left(Y \mid X\right),\varepsilon} B_Y^T q \\ &\Rightarrow 0 &= &\| B_Y^T q\| \Rightarrow q^T B_Y = 0 \end{split}$$

since $\Sigma_{(Y|X)\epsilon}$, is positive definite. But then,

$$\begin{split} \lambda q^T = q^T \left(A - (B_X, B_Y) \left(\sum_{\sum_{YX,\varepsilon}} \sum_{X,\varepsilon} \right) \sum_{X,\varepsilon}^{-1} C_X \right) \\ = q^T \left(A - \left(B_X C_X + B_Y \sum_{YX,\varepsilon} \sum_{X,\varepsilon}^{-1} C_X \right) \\ = q^T \left(A - B_X C_X \right). \end{split}$$

Thus, $(A - B_X C_X, B_Y)$ is not controllable. But this is a contradiction since we can find a matrix, namely C_Y , so that $A - B_X C_X - B_Y C_Y = A - BC$ is stable.

Remarks

i. For implementation in Matlab, positive definiteness in constructing Q can be an issue. A simple resolution is to use a Cholesky factorization of $\sum_{\varepsilon} = L_{\varepsilon} L_{\varepsilon}^{T}$ and form $B_{\varepsilon} = BL_{\varepsilon}$ and then form $Q = B_{\varepsilon} B_{\varepsilon}^{T}$.

ii. The $P_{(X)}$ matrix from DARE,

$$\begin{bmatrix} P_{(X)}, L_0, G \end{bmatrix} = DARE \left(A^T, C_X^T, B \sum_{\varepsilon} B^T, \sum_{X, \varepsilon}, B_o, I \right) \text{ obeys}$$

$$P_{(X)} = AP_{(X)}A^T + B \sum_{\varepsilon} B^T - K_{(X)}\Omega_X K_{(X)}^T \text{ and then}$$

$$K_{(X)} = \left(AP_{(X)}C_X^T + B_o \right) \Omega_X^{-1}, \Omega_X = \sum_{X, \varepsilon} + C_X P_{(X)}C_X^T.$$

iii. Dufour and Taamouti (2010) discuss a method for obtaining sub-models, but it is flawed. First, it requires computing the inverse of the VAR operator. While this might be feasible (analytically) on a toy example, there is no known numerically reliable way to do this in general (computation of determinants is notoriously ill conditioned). Second, it requires the solution of simultaneous quadratic autocovariance equations to determine VMA parameters for which no algorithm is given. In fact, these are precisely the equations required for a spectral

Neural Comput. Author manuscript; available in PMC 2017 August 26.

Solo

factorization of a VMA process. There are reliable algorithms for doing this, but given the flaw already revealed, we need not discuss this approach any further.

Next we state an important corollary:

Corollary 1—Any submodel is in general a VARMA model, not a VAR. To put it another way, the class of VARMA models is closed under the forming of submodels, whereas the class of VAR models is not.

This means that VAR models are not generic and is a strong argument against their use. Any vector time series can be regarded as a submodel of a larger-dimensional time series and thus must in general obey a VARMA model. This result (which is well known in time series folklore) is significant for econometrics where VAR models are in widespread use.

For the next section we need:

Theorem 2—For the joint ISS model, equation 2.3 or 2.4, for z_t with conditions St, De holding and with induced submodel for x_t given in theorem 1, we have,

$$\begin{split} f_{X}\left(\lambda\right) = & h_{X}\left(\exp\left(-j\lambda\right)\right)\Omega_{X}h_{X}^{T}\left(\exp\left(j\lambda\right)\right),\\ h_{X}\left(L\right) = & I + C_{X}\left(L^{-1}I - A\right)^{-1}K_{(X)} = & I + LC_{X}\left(I - AL\right)^{-1}K_{(X)},\\ & ln\left|\Omega_{X}\right| = \int_{-\pi}^{\pi} ln\left|f_{X}\left(\lambda\right)\right|\frac{d\lambda}{2\pi}. \end{split} \tag{3.1}$$

4 Downsampling

There are two approaches to the problem of finding the model obeyed by a downsampled process: frequency domain and time domain. While the general formula for the spectrum of a downsampled process has long been known, it is not straightforward to use and has not yielded any general computational approach to finding submodels of parameterized spectra.

The spectral formula, however, has made it very clear that downsampling leads to aliasing, so that higher-frequency information is lost. This strongly suggests that GCS will be affected by downsampling.

The most complete (time domain) work on downsampling seems to be that of Pandit and Wu (1983), although they treat only the first- and second-order scalar cases. There is work in the engineering literature for systems with observed inputs, but that is also limited and in any case not helpful here. We follow an SS route.

We begin with the ISS model, equation 2.3. Suppose we downsample the observed signal z_t with sampling multiple *m*. Let *t* denote the fine timescale and *k* the coarse timescale so t = mk. The downsampled signal is $\overline{z}_k = z_{mt}$. To develop the SS model for \overline{z}_k , we iterate the SS model above to obtain

$$\xi_{t+l} = A^l \xi_t + \sum_{1}^l A^{l-i} B \varepsilon_{t+i-1}.$$

Now set t = mk, l = m and denote sampled signals, $\overline{\xi}_k = \xi_{mk}, \overline{z}_k = \varepsilon_{mk}$. Then we find

$$\overline{\xi}_{k+1} = A^m \overline{\xi}_k + w_k, \ \overline{z}_k = C \overline{\xi}_k + \overline{\varepsilon}_k,$$

where $w_k = \sum_{1}^{m} A^{m-i} B \varepsilon_{km+i-1}$. We now use result 1 to find the ISS model corresponding to this SS model.

We first have to calculate the model covariances:

$$E(\overline{\varepsilon}_k \overline{\varepsilon}_k^T) = \sum_{\varepsilon} = R,$$

$$E(w_k \overline{\varepsilon}_k^T) = A^{m-1} B \sum_{\varepsilon} = S_m, \quad (4.1)$$

$$E(w_k w_k^T) = Q_m,$$

$$Q_m = \sum_{1}^{m} A^{m-i} B \sum_{\varepsilon} B^T (A^T)^{m-i} = \sum_{0}^{m-1} A^r B \sum_{\varepsilon} B^T (A^T)^r,$$

$$\Rightarrow Q_m = A Q_{m-1} A^T + B \sum_{\varepsilon} B^T, m \ge 2, Q_1 = B \sum_{\varepsilon} B^T,$$
(4.2)

We now obtain:

Theorem 3

Given the ISS model, equation 2.3, then under condition Ev, for m > 1, the ISS model for the downsampled process $\overline{z}_k = z_{mt} is(A^m, C, K_m^*, V_m^*)$, obtained by solving the DARE with SS model (A^m , C, $[Q_m, R, S_m]$) where Q_m is given in equation 4.2 and S_m is given in equation 4.1.

Proof

Using result 1, we need to show the following. R is positive definite, (A^m, C) is detectable,

and $(A^m - S_m R^{-1} C)$, $(Q_m - S_m R^{-1} S_m^T)^{\frac{1}{2}}$ is stabilizable. In fact, we show controllability. The first holds trivially; the second does as well since A is stable and thus so is A^m . For the third, we use the PBH test.

Suppose controllability fails. Then there is a left eigenvector q (possibly complex) with $\lambda q^T = q^T (A^m - S_m R^{-1} C) = q^T A^{m-1} (A - BC)$ and

$$q^{T}(Q_{m}-S_{m}R^{-1}S_{m}^{T})=0=q^{T}\sum_{0}^{m-2}A^{r}B\sum_{\varepsilon}B^{T}(A^{T})^{r}\Rightarrow\sum_{0}^{m-2}q^{H}\sum_{0}^{m-1}A^{r}B\sum_{\varepsilon}B^{T}(A^{T})^{r}q=0$$

Since Σ_{e} is positive definite, this delivers $\sum_{0}^{m-2}||B^{T}(A^{T})^{r}q||^{2}=0\Rightarrow B^{T}(A^{T})^{r}q=0$ for $r=0, ..., m-2$.

Using this, we now find $\lambda - q^T = q^T A^{m-1}(A - BC) = q^T A^{m-2}(A - BC)^2 + q^T A^{m-2}BC(A - BC) = q^T A^{m-2}(A - BC)^2$. Iterating this yields $\lambda q^T - q^T(A - BC)^m$. Thus, if λ_m is an *m*th

root of λ , then $\lambda_m q^T = q^T (A - BC)$. Since $q^T B = 0$, we thus conclude that (A - BC, B) is not controllable. But this is a contradiction since (A - BC) + BC = A is stable.

Remark

i. In Matlab, we would compute,

$$[P_m^*, L_0, G_m] = \text{DARE}((A^m)^T C_T, Q_m, \sum_{\varepsilon}, S_m, I)$$
, yielding $V_m^* = \sum_{\varepsilon} + CP_m^* C^T$
and $K_m^* = G_m^T$.

ii. More specifically
$$P_m^*(m>1)$$
 obeys $P_m^* = \overline{A}_m P_m^* \overline{A}_m^T + \overline{Q}_m - \overline{K}_m V_m^* \overline{K}_m^T$, where
 $\overline{A}_m = A^m - S_m \sum_{\varepsilon}^{-1} C = A^{m-1} (A - BC); \overline{Q}_m = Q_m - S_m \sum_{\varepsilon}^{-1} S_m^T = Q_{m-1};$
 $V_m^* = \sum_{\varepsilon} + C P_m^* C^T; K_m^* = \overline{A}_m P_m^* C^T (V_m^*)^{-1}.$

5 Granger Causality

In this section, we review and extend some basic results in Granger causality. In particular, we extend GEMs to the state-space setting and show how to compute them reliably.

Since the development of Granger causality, it has become clear (Dufour & Renault, 1998; Dufour & Taamouti, 2010) that in general, one cannot address the causality issue with only one-step-ahead measures as commonly used. One needs to look at causality over all forecast horizons. However one-step measures are sufficient when one is considering only two vector time series as we are (Dufour & Renault, 1998, proposition 2.3).

5.1 Granger Causality Definitions

Our definitions of one-step Granger causality naturally draw on Granger (1963, 1969), Sims (1972), and Solo (1986) but are also influenced by Caines (1976), who, drawing on the work of Pierce and Haugh (1977), distinguished between weak and strong GC, or what Caines calls weak and strong feedback free processes. We introduce:

- Condition WSS: The vector time series x_t , y_t are jointly second-order stationary.
- Definition: Weak Granger causality. Under WSS, we say y_t does not weakly Granger-cause (dn-wgc) x_t if, for all t,

$$var\left(X_{t+1}|X_{-\infty}^{t},Y_{-\infty}^{t}\right) = var\left(X_{t+1}|X_{-\infty}^{t}\right).$$

Otherwise, we say y_t weakly Granger causes (wgc) x_t . Because of the elementary identity,

$$var(X|Z) = E \left[var(X|Z,W) \right] + var \left[E(X|Z,W) \right]$$
$$= E \left[var(X|Z,W) \right] + E \left[(E(X|Z,W) - E(X|Z))^T |Z \right],$$

the equality of variance matrices in the definition also ensures the equality of

predictions: $E\left(X_{t+1}|X_{-\infty}^{t},Y_{-\infty}^{t}\right) = E\left(X_{t+1}|X_{-\infty}^{t}\right)$. This definition agrees with Granger (1969); Caines and Chan (1975), who do not use the designator weak; and Caines (1976) and Solo (1986) who do.

• Definition: Strong Granger causality. Under WSS, we say y_t does not strongly Granger cause (dn-sgc) x_t if, for all t,

$$var(X_{t+1} | X_{-\infty}^t, Y_{-\infty}^t, Y_{t+1}) = var(X_{t+1} | X_{-\infty}^t).$$

Otherwise we say y_t strongly Granger-causes (sgc) x_t . Again, equality of the variance matrices ensures equality of predictions,

 $E(X_{t+1} | X_{-\infty}^t, Y_{-\infty}^t, Y_{t+1}) = E(X_{t+1} | X_{\infty}^t)$. This definition agrees with Caines (1976) and Solo (1986).

- Definition: FBI (feedback interconnected). If x_t Granger-causes y_t and y_t Granger-causes x_t , then we say x_t , y_t are feedback interconnected.
- Definition: UGC (unidirectionally Granger causes). If x_t Granger-causes y_t but y_t dn-gc x_t we say x_t unidirectionally Granger-causes y_t .

5.2 Granger Causality for Stationary State-Space Models

Now we partition $z_t = (x_b \ y_t)^T$ into subvector signals of interest and partition the vector MA or state-space model, equation 2.4, correspondingly:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \left[I + \begin{bmatrix} C_X \\ C_Y \end{bmatrix} \left(L^{-1}I - A\right)^{-1} \left(B_X B_Y\right) \begin{bmatrix} \varepsilon_{X,t} \\ \varepsilon_{Y,t} \end{bmatrix}$$
(5.1)

$$= \begin{bmatrix} H_{XX}(L)H_{XY}(L) \\ H_{YX}(L)H_{YY}(L) \end{bmatrix} \begin{bmatrix} \varepsilon_{X,t} \\ \varepsilon_{Y,t} \end{bmatrix}$$
$$\sum_{\varepsilon} = var \begin{bmatrix} \varepsilon_{X,t} \\ \varepsilon_{Y,t} \end{bmatrix} = \begin{bmatrix} \sum_{X,\varepsilon} & \sum_{XY,\varepsilon} \\ \sum_{YX,\varepsilon} & \sum_{Y,\varepsilon} \end{bmatrix}$$
$$\begin{bmatrix} H_{XX}(L)H_{XY}(L) \\ H_{YX}(L)H_{YY}(L) \end{bmatrix} = \begin{bmatrix} C_{X}(L^{-1}I - A)^{-1}B_{X} + I \ C_{X}(L^{-1}I - A)^{-1}B_{Y} \\ C_{Y}(L^{-1}I - A)^{-1}B_{X} \ C_{Y}(L^{-1}I - A)^{-1}B_{Y} + I \end{bmatrix}.$$
(5.2)

Now we recall results of Caines (1976):

Result 3—If $z_t = \begin{pmatrix} x_t \\ y_t \end{pmatrix}$ obeys a Wold model of the form $z_t = H_Z(L)\varepsilon_t$ where $H_Z(L)$ is a one-sided square summable moving average polynomial with $H_Z(0) = I$, which is partitioned as in equation 5.2, then:

- **a.** y_t dn-wgc x_t iff $H_{XY}(L) = 0$.
- **b.** $y_t \operatorname{dn-sgc} x_t \operatorname{iff} H_{XY}(L) = 0 \text{ and } \Sigma_{XY,e} = 0.$

We can now state a new SS version of this result:

Theorem 4—For the stationary ISS model, equations 5.1 and 5.2:

- **a.** y_t dn-wgc x_t iff $C_X A^r B_Y = 0, r = 0$.
- **b.** y_t dn-sgc x_t iff $C_X A^r B_Y = 0$, r = 0 and $\Sigma_{XY,\epsilon} = 0$.

Proof—The proof follows immediately from result 3 since $H_{XY}(L) = \sum_{0}^{\infty} C_X A^r B_Y L^{r+1}$.

Remarks

- i. By the Cayley-Hamilton theorem, we can replace part a of theorem 4 with $C_X A^r B_Y = 0, 0 \quad r \quad n-1, n = \dim(\xi_l).$
- **ii.** Collecting these equations together gives $C_X(B_Y, AB_Y, ..., A^{n-1}B_Y) = 0$, which says that the pair (A, B_Y) is not controllable. Also, we have

 $B_Y^T(C_X^T, A^T C_X^T, \dots, (A^T)^{n-1} C_X^T) = 0$, which says that the pair (C_X , A) is not observable. Thus, the representation of $H_{XY}(L)$ is not minimal.

From a data analysis point of view, we need to embed this result in a well-behaved hypothesis test. The results of Geweke (1982), suitably modified, allow us to do this.

5.3 Geweke Causality Measures for SS Models

Although much of the discussion in Geweke (1982) is in terms of VARs, we can show it applies more generally. We begin as Geweke (1982) did with the following definitions. First,

$$F_{Y \to X} = ln \frac{|\Omega_X|}{\left|\sum_{x,\varepsilon}\right|}$$
 is a measure of the gain in using the past of y to predict x beyond using

just the past of *x*; similarly introduce $F_{Y \to X} = ln \frac{|\Omega_Y|}{|\sum_{Y, \in}|}$. Thus, for example, $F_{Y \to X}$ is a directional measure of the influence of *yt* on *x_t*. Next, define the instantaneous influence

measure: $F_{Y,X} = ln \frac{|\sum_{X,\varepsilon} ||\sum_{Y,\varepsilon}|}{|\sum_{\varepsilon}|}$. These are then joined in the fundamental decomposition (Geweke, 1982),

$$F_{X \circ Y} = F_{Y \to X} + F_{X \to Y} + F_{Y.X}$$
 (5.3)

where, $F_{x_{oY}} = ln \frac{|\Omega_x| |\Omega_y|}{|\sum_{\varepsilon}|}$. Geweke (1982) then proceeds to decompose these measures in the frequency domain. Thus, the frequency domain GEM for the dynamic influence of y_t on x_t is given by Geweke (1982),

$$F_{Y \to X} = \int_{-\pi}^{\pi} f_{Y \to X}(\lambda) \frac{d\lambda}{2\pi} \text{where } f_{Y \to X}(\lambda) = \ln \frac{|f_X(\lambda)|}{|f_e(\lambda)|}, \quad (5.4)$$

and $f_c(\lambda)$ is assembled (following Geweke, 1982) as follows.

Introduce $W = \sum_{YX,\varepsilon} \sum_{X,\varepsilon}^{-1}$ and note that $\varepsilon_{Y,t} - W \varepsilon_{X,t}$ is uncorrelated with $\boldsymbol{e}_{X,t}$ and has variance $\sum_{(Y|X),\varepsilon} = \sum_{Y,\varepsilon} -\sum_{YX,\varepsilon} \sum_{X,\varepsilon}^{-1} \sum_{XY,\varepsilon}$. Then rewrite equation 5.2 as $\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} H_{XX}(L)H_{XY}(L) \\ H_{YX}(L)H_{YY}(L) \end{bmatrix} \begin{bmatrix} I & 0 \\ W & I \end{bmatrix} \begin{bmatrix} I & 0 \\ -W & I \end{bmatrix} \begin{bmatrix} \varepsilon_{X,t} \\ \varepsilon_{Y,t} \end{bmatrix}$ $= \begin{bmatrix} H_{XX}(L)+H_{XY}(L)W, H_{XY}(L) \\ H_{YX}(L)+H_{YY}(L)W, H_{YY}(L) \end{bmatrix} \begin{bmatrix} \varepsilon_{X,t} \\ \varepsilon_{Y,t} \end{bmatrix}$.

This corresponds to equation 3.3 in Geweke (1982) and yields the following expressions corresponding to those in Geweke (1982):

$$f_{X}(L) = f_{e}(L) + H_{XY}(L) \sum_{(Y|X),\varepsilon} H_{XY}^{T}(L^{-1}),$$
(5.5)

$$f_{e}(L) = H_{eX}(L) \sum_{X,\varepsilon} H_{eX}^{T}(L^{-1}), H_{eX}(L) = H_{XX}(L) + H_{XY}(L)W.$$
(5.6)

Using the SS expressions above, we rewrite $H_{eX}(L)$ in a form more suited to computation as

$$H_{e_X}(L) = [C_X (L^{-1}I - A)^{-1}B_o + I]$$

$$B_o = B_X + B_Y \sum_{XY,\varepsilon}^T \sum_{X,\varepsilon}^{-1} = B\left(\sum_{YX,\varepsilon}^{X,\varepsilon}\right) \sum_{X,\varepsilon}^{-1}.$$

(5.7)

Note that then, using theorem 2,

$$F_{Y \to X} = \int_{-\pi}^{\pi} ln |f_X(\lambda)| \frac{d\lambda}{2\pi} - \int_{-\pi}^{\pi} ln |f_e(\lambda)| \frac{d\lambda}{2\pi}$$
$$= ln |\Omega_X| - ln |\sum_{X,\varepsilon}| = ln \frac{|\Omega_X|}{|\sum_{X,\varepsilon}|}.$$
(5.8)

Clearly, with $L = \exp(-j\lambda)$, $f_X(L) \quad f_e(L) \Rightarrow F_{Y \to X} = 0$.

Also the instantaneous causality measure is

$$F_{Y,X} = ln \frac{|\sum_{X,\varepsilon}||\sum_{Y,\varepsilon}|}{|\sum_{\varepsilon}|} = ln \frac{|\sum_{X,\varepsilon}||\sum_{Y,\varepsilon}|}{|\sum_{(X|Y),\varepsilon}||\sum_{Y,\varepsilon}|} = ln \frac{|\sum_{X,\varepsilon}|}{\sum_{(X|Y),\varepsilon}|}$$
$$\sum_{(X|Y),\varepsilon} = \sum_{X,\varepsilon} - \sum_{XY,\varepsilon} \sum_{Y,\varepsilon}^{-1} \sum_{YX,\varepsilon} .$$
(5.9)

Clearly $\Sigma_{X,e}$ $\Sigma_{(X|Y),e}$ so that $F_{Y,X}$ 0.

Introduce the normalized cross-covariance-based matrix,

 $\Gamma_{x,y} = \sum_{Y,\varepsilon}^{-\frac{1}{2}} \sum_{YX,\varepsilon} \sum_{X,\varepsilon}^{-1} \sum_{YX,\varepsilon} \sum_{Y,\varepsilon}^{-\frac{1}{2}}$. Then, using a well-known partitioned matrix determinant formula (Magnus & Neudecker, 1999), we find $F_{YX} = In|I - \Gamma_{x,y}|$. This means that the instantaneous causality measure depends only on the canonical correlations (which are the eigenvalues of $\Gamma_{x,y}$) between $\varepsilon_{X,t}$, $\varepsilon_{Y,t}$ (Seber, 1984; Kshirsagar, 1972).

To implement these formulas, we need expressions for Ω_X , Ω_Y , $f_X(\lambda)$. To get them, Geweke (1982) fits separate models to each of x_t and y_t . But this causes positivity problems with $f_{Y \to X}(\lambda)$ (Chen et al., 2006). Instead we obtain the required quantities from the correct submodel obtained in the previous section. We have:

Theorem 5a—The GEMs can be obtained from the joint ISS model, equation 5.1, and the submodel in theorem 2, as follows:

- **a.** $F_{Y \to X} = ln \frac{|\Omega_X|}{|\sum_{X,\varepsilon}|}$ where Ω_X is got from the submodel in theorem 3.
- **b.** The frequency domain GEM $f_{Y\to X}(\lambda)$, equation 5.4, can be computed from equations 5.6, 3.1, and 5.7.

 $\Omega_Y, F_{X \to Y}, f_{X \to Y}(\lambda)$ can be obtained similarly.

Now pulling all this together with the help of result 3, we have an extension of the results of Geweke (1982) to the state-space/VARMA case.

Theorem 5b—For the joint ISS model, equation 5.1:

a.

$$F_{Y \to X} + F_{Y,X} = ln \frac{|\Omega_X|}{\left| \sum_{(X|Y), \varepsilon} \right|}$$

F_Y \to X 0, F_{Y,X} 0 and

- **b.** y_t dn-wgc x_t iff, $f_X(L) = f_e(L)$ which holds iff $F_{Y \to X} = 0$, that is, iff $\Omega_X = \Sigma_{X,e}$.
- **c.** $y_t \text{ dn-sgc } x_t \text{ iff } f_X(L) = f_e(L) \text{ and } \Sigma_{(X|Y),\epsilon} = \Sigma_{X,\epsilon} \text{ that is, iff } F_{Y \to X} = 0 \text{ and}$ $F_{Y,X}=0, \text{ that is, iff } F_{Y \to X} + F_{Y,X} = 0 \text{ i.e. iff } \Omega_X = \Sigma_{(X|Y),\epsilon}.$

Remarks

- **i.** A very nice nested hypothesis testing explanation of the decomposition 5.3 is given by Parzen in the discussion to Geweke (1982).
- **ii.** It is straightforward to see that the GEMs are unaffected by scaling of the variables. This is a problem for other GC measures (Edin, 2010).
- iii. For completeness, we state extensions of the inferential results in Geweke (1982) without proof. Suppose we fit an SS model to data z_b t = 1, ..., T using, for example, so-called state-space subspace methods (Deistler et al., 1995; Bauer,

2005) or VARMA methods in, for example, Lutkepohl (1993). Let $\hat{F}_{Y \to X}, \hat{F}_{X \to Y}, \hat{F}_{Y,X}, \hat{F}_{X o Y}$ be the corresponding GEM estimators. If we denote true values with

a superscript 0, we find under some regularity conditions:

$$\begin{split} H_0: F^0_{Y \to X} = & 0 \Rightarrow T\hat{F}_{Y \to X} \Rightarrow \chi^2_{2np_x p_Y}, \text{ as } T \to \infty \\ & \text{and } H_0: F^0_{Y,X} = & 0 \Rightarrow T\hat{F}_{Y,X} \Rightarrow \chi^2_{p_x p_Y}. \end{split}$$

So to test for strong GC, we put these together:

$$H_0: F^0_{Y \to X} = 0, F^0_{Y,X} = 0 \Rightarrow T(\hat{F}_{Y \to X} + \hat{F}_{Y,X}) \Rightarrow \chi^2_{(2n+1)p_x p_Y}.$$

Together with similar asymptotics for $\hat{F}_{X \to Y}$, $\hat{F}_{X \circ Y}$, we see that the fundamental decomposition, equation 5.3, has a sample version involving a decomposition of a chi-squared into sums of smaller chi-squared statistics.

- iv. Chen et al. (2006) also attempt to derive $F_{Y \to X}$ without fitting separate models to x_b, y_t . However the proposed procedure to compute $f_X(\lambda)$ involves a two-sided filter and is thus in error. The only way to get $f_X(\lambda)$ is by spectral factorization (which produces one-sided or causal filters), as we have done.
- v. Other kinds of causality measures have emerged in the literature (Kaminski & Blinowska, 1991; Baccala & Sameshima, 2001; Takahashi, Baccala, & Sameshima, 2008; Pascual-Marqui et al., 2014), but it is not known whether they obey the properties in theorems 5a and 5b. However these properties are crucial to our subsequent analysis.

6 Effect of Filtering on Granger Causality Measures

Now the import of the frequency domain GEM becomes apparent since it allows us to determine the effect of one-sided (or causal) filtering on GC.

We need to be clear on the situation envisaged here. The unfiltered time series are the underlying series of interest, but we have access only to the filtered time series, so we can only find the GEMs from the spectrum of the filtered time series. What we need to know is when those filtered GEMs are the same as the underlying GEMs. We have:

Theorem 6

Suppose we filter z_t with a stable, full-rank, one-sided filter $\Phi(L) = \begin{pmatrix} \Phi_{_{XX}}(L) & 0 \\ 0 & \Phi_{_{YY}}(L) \end{pmatrix}$. Then:

- **a.** If $\Phi(L)$ is minimum phase, the GEMs (and so GC) are unaffected by filtering.
- **b.** If $\Phi(L)$ has the form $\Phi(L) = \psi(L)\overline{\Phi}(L)$ where $\psi(L)$ is a scalar all-pass filter and $\overline{\Phi}(L)$ is stable, minimum phase, then the GEMs (and so GC) are unaffected by filtering.
- c. If $\Phi(L)$ is nonminimum phase and case b does not hold, then the GEMs (and so GC) are changed by filtering.

Proof

Denote $\overline{z}_t = \Phi(L)z_t = \Phi(L)H(L)\varepsilon_t$ by result 2a. Then for the frequency domain GEM, we

need to find $f_{\overline{Y}\to\overline{X}}(\lambda) = ln \frac{|f_{\overline{X}}(\lambda)|}{|H_{e\overline{X}}(L)\sum_{X,e}H_{e\overline{X}}^{T}(L^{-1})|}$, where $L = \exp(-j\lambda)$. We find trivially that $|f_{\overline{X}}(\lambda)| = |\Phi_{X}(L)f_{X}(\lambda)\Phi_{X}(L^{-1})| = |\Phi_{X}(L)||f_{X}(\lambda)||\Phi_{X}(L^{-1})|$. Finding $H_{e\overline{X}}(L)$ is much more complicated; we need the minimum phase vector moving average or state-space model corresponding to equation 5.2. Taking $\Phi(L)$ to be nonminimum phase, we carry out a spectral factorization, $f_{\overline{z}}(\lambda) = \overline{H}(L)\overline{\sum}\overline{H}^{T}(L^{-1})$, where $\overline{H}(L)$ is causal, stable, minimum phase with $\overline{H}(0) = I$ and then, from appendix C, $\overline{H}(L)$, can be written $\overline{H}(L) = \Phi(L)H(L)D(L^{-1}), D(L^{-1}) = JE^{T}(L^{-1})\overline{J}^{-1}$, where E(L) is all pass and J, \overline{J} are

 $H(L) = \Phi(L)H(L)D(L^{-1}), D(L^{-1}) = JE^{-1}(L^{-1})J^{-1}$, where E(L) is all pass and J, \overline{J} are constant matrices (Cholesky factors). Writing this in partitioned form,

$$\overline{H}(L) = \begin{bmatrix} \Phi_{X}(L) & 0\\ 0 & \Phi_{Y}(L) \end{bmatrix} \begin{bmatrix} H_{XX}(L) & H_{XY}(L)\\ H_{YX}(L) & H_{YY}(L) \end{bmatrix} \begin{bmatrix} D_{XX}(L^{-1}) & D_{XY}(L^{-1})\\ D_{YX}(L^{-1}) & D_{YY}(L^{-1}) \end{bmatrix}$$

yields $H_{e\overline{X}}(L) = \Phi_X(L) K_{e\overline{X}}(L)$, where

$$\begin{split} &K_{e\overline{X}}(L) {=} H_{\scriptscriptstyle XX}(L) (D_{\scriptscriptstyle XX}(L^{-1}) {+} D_{\scriptscriptstyle XY}(L^{-1}) {\textstyle \sum_{YX,\varepsilon} \sum_{X,\varepsilon}^{-1}} \\ &+ H_{\scriptscriptstyle XY}(L) (D_{\scriptscriptstyle YY}(L^{-1}) {\textstyle \sum_{YX,\varepsilon} \sum_{X,\varepsilon}^{-1}} {+} D_{\scriptscriptstyle YX}(L^{-1})). \end{split}$$

Thus, in $f_{\overline{Y} \to \overline{X}}(\lambda)$, the $|\Phi_X(L)|$ factors cancel giving,

$$\begin{split} f_{\overline{Y} \to \overline{X}}\left(\lambda\right) = & ln \frac{|f_X(\lambda)|}{\left|K_{e\overline{X}}\left(L\right) \sum_{X,\varepsilon} K_{e\overline{X}}^T \left(L^{-1}\right)\right|}. \text{ This will reduce to } f_{Y \to X}(\lambda) \text{ iff} \\ K_{e\overline{X}}(L) = & H_{eX}(L)K(L^{-1}) \text{ where } K(L^{-1}) \text{ is all pass, which occurs iff } D_{XY}(L^{-1}) = 0, \\ D_{YX}(L^{-1}) = & 0, D_{XX}(L^{-1}) = \psi(L^{-1})I, D_{YY}(L^{-1}) = \psi(L^{-1})I \text{ where } \psi(L^{-1}) \text{ is a scalar all-pass} \\ \text{filter. Results a, b, and c now follow.} \end{split}$$

Remark

Result a has also been obtained in the independent work of Seth, Chorley, and Barnett (2013) by different methods.¹

We give three examples:

Example 6A

Differential delay. Suppose
$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \rho & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_t \\ \nu_t \end{pmatrix}$$
 where \boldsymbol{e}_t , \boldsymbol{v}_t are independent zero
mean white poises with variances $1 - \sigma^2$ respectively, while $\Phi(L) = \begin{pmatrix} 1 & 0 \\ 0 & L \end{pmatrix}$. So the two

mean white noises with variances 1, σ^2 , respectively, while $\begin{pmatrix} 0 & L \end{pmatrix}$. So the two series are white noises that exhibit an instantaneous GC. The filtering delays one series relative to the other. Then we have

$$\overline{z}_t = \begin{pmatrix} 1 & 0 \\ 0 & L \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \rho & 1 \end{pmatrix} \begin{pmatrix} a_t \\ b_t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ L_\rho & 1 \end{pmatrix} \begin{pmatrix} a_t \\ b_t \end{pmatrix} = \overline{H}(L) \begin{pmatrix} a_t \\ b_t \end{pmatrix}, \text{ and we see that}$$

 $\overline{H}(0) = I$ while $\overline{H}(L)$ is stable, causal, and invertible; indeed, $1 - L\rho - 1$. Thus, we see that the differential delay has introduced a spurious dynamic GC relation and the original purely instantaneous GC is lost.

Example 6B

Noncausal filter. We use the same setup as in example 6 A except that now

$$\Phi(L) = \begin{pmatrix} 1 - \theta_x L^{-1} & 0 \\ 0 & 1 - \theta_y L^{-1} \end{pmatrix} \text{ where } |\theta_x| < 1, |\theta_y| < 1 \text{ and } \theta_x \quad \theta_y. \text{ We rewrite this as } \\ -L^{-1}\Psi(L) \text{ where } \Psi(L) = \begin{pmatrix} \theta_x (1 - \frac{1}{\theta_x}L) & 0 \\ 0 & \theta_y (1 - \frac{1}{\theta_y}L) \end{pmatrix}. \text{ Although } L^{-1} \text{ is noncausal and so NMP, it is the same for both filters and so by theorem 6b will not affect the GEMS. So we can replace $\Phi(L)$ with $\Psi(L)$. Both entries in $\Psi(L)$ are causal but unstable, and so NMP and the GCS will be altered. In fact, we now show this explicitly.$$

Note, first, how the noncausal filter operates. We have

$$\overline{z}_t {=} \Psi(L) z_t {=} \left(\begin{array}{cc} \theta_x - L & 0 \\ 0 & \theta_y - L \end{array} \right) \left(\begin{array}{c} x_t \\ y_t \end{array} \right) {=} \left(\begin{array}{c} \theta_x x_t - x_{t+1} \\ \theta_y y_t - y_{t+1} \end{array} \right).$$

Future values are needed to compute the filtered signal at the current time.

¹The main parts of this theorem were announced in a research workshop at the 2011 Human Brain Mapping meeting

Neural Comput. Author manuscript; available in PMC 2017 August 26.

 $\text{The process } \overline{z}_t = \Psi(L) z_t \text{ has spectrum} \begin{pmatrix} AA* & \rho A*B\\ \rho B*A & (\rho^2 + \sigma^2)BB* \end{pmatrix}_{\text{where}}$

 $A=A(L)=\theta_x\left(1-\frac{1}{\theta_x}L\right)$, $A*=A(L^{-1})$ and similarly for *B*. It is straightforward to check that this has spectral factorization,

$$\begin{pmatrix} a & \frac{\rho}{R^2}a \\ 0 & b \end{pmatrix} \begin{pmatrix} \frac{\alpha^2}{R^2} & 0 \\ 0 & R^2 \end{pmatrix} \begin{pmatrix} a* & 0 \\ \frac{\rho}{R^2}a* & b* \end{pmatrix} = \begin{pmatrix} aa* & \rho b*a \\ \rho ba* & R^2bb* \end{pmatrix}$$

where $R = \sqrt{\rho^2 + \sigma^2}$ and $a = 1 - \theta_x L$, $b = 1 - \theta_y L$. Note that $AA^* = aa^*$ and $A^*B = ab^*$. This

factorization corresponds to the model $\begin{pmatrix} \overline{x}_t \\ \overline{y}_t \end{pmatrix} = \begin{pmatrix} a & \frac{\rho}{R^2}a \\ 0 & b \end{pmatrix} \begin{pmatrix} \varepsilon_{a,t} \\ \nu_{b,t} \end{pmatrix}$ where $\varepsilon_{a,b}$, $v_{b,t}$ are

independent white noises of variances $\frac{\rho^2}{R^2}$, R^2 , respectively. We see that the NMP filtering has introduced spurious dynamics and reversed the GCS.

Example 6C

Wiener deconvolution. When filtered versions of x_t , y_t are observed in noise, one can estimate them by using the Wiener fitler (Kailath et al., 2000). The optimal filter uses the joint spectrum, but fMRI practice (Glover, 1999) is to filter each signal separately. This is suboptimal but will still reduce the effect of the noise. Suppose we observe $w_{x,t} = h_x(L)x_t + e_t$ where h(L) is a stable filter and e_t is a white noise of variance σ^2 independent of x_t . The Wiener filter estimate of x_t is $\hat{x}_t = W_x(L)w_{x,t}$ where (Kailath et al., 2000)

 $W_x(L) = \frac{h_x^*}{h_x h_x^* + \sigma^2 / \gamma_x} \text{ where } \gamma_x = \gamma_x(L) \text{ is the autocovariance generating function of } x_t \text{ In fMRI practice (Glover, 1999), } \gamma_x(L) \text{ is set to } 1. \text{ This is sub-optimal and may lead to performance worse than no filtering. But continuing, we carry out a spectral factorization giving } h_x h_x^* + \sigma^2 = \sigma_e^2 k_x k_x^* \text{ where } k_x = k_x(L) \text{ is a causal minimum phase filter with } k_x(0) = 1.$ The filter thus becomes $W_x(L) = \frac{h_x^*}{\sigma_e^2 k_x k_x^*}$. Since k_x is minimum phase and causal, then k_x^{-1} is

stable and causal and will not affect the GEMS. We are left with the noncausal filter $\beta_x^* = \frac{h_x^*}{k_x^*}$ — for example, in the simple case, $h_x = 1 - \theta_x L$, $|\theta_x| < 1$, the spectral factorization reduces to solving a quadratic equation and one obtains $k_x = 1 - a_x L$, $|a_x| < 1$ Thus,

 $\beta_x^* = \frac{1 - \theta_x L^{-1}}{1 - \alpha_x L^{-1}} = 1 + \sum_{1}^{\infty} (\alpha_x - \theta_x) \alpha_x^{r-1} L^{-r}.$ With a similar deconvolution (where $\beta_x \beta_y$) applied to a noisy filtered version of y_b we conclude that deconvolution will distort the GCS since the β_x , β_y filters are noncausal.

7 Downsampling and Forwards Granger Causality

We now consider to what extent GC is preserved under downsampling.

Using the sampled notation of our discussion above and defining $\overline{z}_x = \begin{pmatrix} \overline{x}_k \\ \overline{y}_k \end{pmatrix}$, we have the following result:

Theorem 7

Forwards causality:

- **a.** If y_t dn-sgc x_t , then \overline{y}_k dn-sgc \overline{x}_k .
- **b.** If y_t dn-wgc x_t , then in general \overline{y}_k wgc \overline{x}_k .

Remarks

- i. Part a is new, although technically a special case of a result of the author's established in a non-SS framework.
- ii. We might consider taking part b as a formalization of longstanding folklore in econometrics (Christiano & Eichenbaum, 1987; Marcet, 1991) that downsampling can destroy unidirectional Granger causality. However, that same folklore is flawed because it failed to recognize the possibility of part a. The folklore is further flawed because it failed to recognize the more serious reverse problem discussed below.

Proof of Theorem 7, Part a—We use the partitioned expressions in the discussion leading up to result 3. We also refer to the discussion leading up to theorem 6.

This allows us to write two decompositions. First, we write $w_k = w_{X,k} + w_{Y,k}$, where

$$w_{X,k} = \sum_{1}^{m} A^{m-i} B_X \varepsilon_{X,km+i-1}, \ w_{Y,k} = \sum_{1}^{m} A^{m-i} B_Y \varepsilon_{X,km+i-1}.$$

From result 3 and the definition of dn-sgc,

$$w_{X,k}$$
 is uncorrelated with $w_{Y,l}$ for all $k, l.$ (7.1)

The other decomposition is $\overline{\varepsilon}_k = \begin{pmatrix} \overline{\varepsilon}_{X,k} \\ \overline{\varepsilon}_{Y,k} \end{pmatrix}$, and

 $\overline{\varepsilon}_{X,k}$ is uncorrelated with $\overline{\varepsilon}_{Y,l}$ for all k, l. (7.2)

Next we note from theorem 4 that

$$C_{X}(L^{-1}I - A^{m})^{-1}A^{p}B_{Y} = C_{X}\sum_{1}^{\infty}A^{m+r-1}B_{Y}L^{r} = 0.$$

for all p = 0. Thus, we deduce

$$C_X w_{Y,k} = 0$$
, for all k. (7.3)

We can now write

$$\begin{aligned} \overline{x}_{k} &= C_{X} (L^{-1}I - A^{m})^{-1} w_{k} + \overline{\varepsilon}_{X,k} \\ &= C_{X} (L^{-1}I - A^{m})^{-1} w_{X,k} + \overline{\varepsilon}_{X,k} \\ \overline{y}_{k} &= C_{Y} (L^{-1}I - A^{m})^{-1} w_{X,k} + C_{Y} (L^{-1}I - A^{m})^{-1} w_{Y,k} + \overline{\varepsilon}_{Y,k}. \end{aligned}$$
(7.4)

Based on equation 7.4, we now introduce the ISS model for $\overline{x}_{k'}$

$$\overline{x}_k = C_X (L^{-1}I - A^m)^{-1} K_{(X)} \nu_{X,k} + \nu_{X,k}$$

where $v_{X,k}$ is the innovations sequence. Using this, we introduce the estimator $K_{(X)}v_{X,k}$ of $w_{X,k}$ and the estimation error $\tilde{w}_{X,k} = w_{X,k} - K_{(X)}v_{X,k}$. Below we show

$$\tilde{w}_{X,k}$$
 is uncorrelated with $\nu_{X,l}$ for all $k, l.$ (7.5)

We thus rewrite the model for \overline{y}_k as

$$\overline{y}_{k} = C_{Y} (L^{-1}I - A^{m})^{-1} K_{(X)} \nu_{X,k} + \zeta_{k},$$

$$\zeta_{k} = C_{Y} (L^{-1}I - A^{m})^{-1} [w_{Y,k} + \tilde{w}_{X,k}] + \overline{\varepsilon}_{Y,k}.$$

Now we can construct an ISS model for $\zeta_k = (I + C_Y(L^{-1} I - A^m)^{-1}K_{(Y)})v_{Y,k}$ where $v_{Y,k}$ is the innovations sequence. In view of equations 7.1, 7.2, and 7.5, $v_{X,k}$ and $v_{Y,l}$ are uncorrelated for all k, l. Thus, we have constructed the joint ISS model:

$$\begin{bmatrix} \overline{x}_k \\ \overline{y}_k \end{bmatrix} = \overline{H}(z) \begin{bmatrix} \nu_{X,k} \\ \nu_{Y,k} \end{bmatrix}$$
$$\overline{H}(z) = \begin{bmatrix} I + C_X (L^{-1}I - A^m)^{-1} K_{(X)} & 0 \\ C_Y (L^{-1}I - A^m)^{-1} K_{(X)} & I + C_Y (L^{-1}I - A^m)^{-1} K_{(Y)} \end{bmatrix}.$$

From this we deduce that \overline{y}_k dn-sgc \overline{x}_k as required.

Proof of Equation 7.5—Consider then

 $E(\tilde{w}_{X,k}\nu_{X,l}^{T}) = E(w_{X,k} - K_{(X)}\nu_{X,k})\nu_{X,l}^{T}) = E(w_{X,k}\nu_{X,l}^{T} - K_{(X)}E(\nu_{X,k}\nu_{X,l}^{T})).$ The second term vanishes for k *I*. The first term vanishes for k > I since $w_{X,k}$ is uncorrelated with the past and hence $v_{X,k}$, for l > k, it vanishes since $v_{X,l}$ is uncorrelated with the past. For k = l, it vanishes by the definition of $K_{(X)}$ (Kailath et al., 2000).

Proof of Theorem 7, Part b—A perusal of the proof of theorem 7, part a, shows that we cannot construct the block lower triangular joint ISS model. In general, we obtain a full block ISS model.

8 Downsampling and Reverse Granger Causality

We now come to the more serious issue of how downsampling might distort GC. To establish distortion, we simply have to exhibit a numerical example, but that is not as simple as one might hope.

8.1 Simulation Design

Designing a procedure to generate a wide class of examples of spurious causality is not as simple as one might hope. We develop such a procedure for a bivariate vector autoregression of order one: a bivariate VAR(1). On the one hand, this is about the simplest example one can consider; on the other hand, it is general enough to generate important behaviors.

The bivariate VAR(1) model is then

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = A \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{X,t} \\ \varepsilon_{Y,t} \end{pmatrix}, A = \begin{pmatrix} \phi_x & \gamma_x \\ \gamma_y & \phi_y \end{pmatrix}, \sum = \begin{pmatrix} \sigma_a^2 & \rho \sigma_a \sigma_b \\ \rho \sigma_a \sigma_b & \sigma_b^2 \end{pmatrix},$$

Where Σ is the variance matrix of the zero mean white noise $\begin{pmatrix} \varepsilon_{X,t} \\ \varepsilon_{Y,t} \end{pmatrix}$ and ρ is a correlation.

We note that this model can be written as an ISS model with parameters, A, I, -A, Σ . Hence, all the computations described above are easily carried out.

But the real issue is how to select the parameters. By a straightforward scaling argument, it is easy to see that we may set $\sigma_a = 1 = \sigma_b$ without loss of generality. Thus, we need to choose only A, ρ .

Some reflection shows that there are two issues. First, we must ensure the process is stationary; for the eigenvalues λ_1 , λ_2 of A, we must have $|\lambda_1| < 1$, $|\lambda_2| < 1$. Second, to design a simulation, we need to choose $F_{Y \to X}$, $F_{X \to Y}$, but these quantities depend on the parameters A, ρ in a highly nonlinear way, so it is not obvious how to do this. And five parameters are already too many to pursue this by trial and error.

For the first issue, we have $trace(A) = \lambda_1 + \lambda_2 = \phi_x + \phi_y$ and $det(A) = \lambda_1 \lambda_2 = \phi_x \phi_y - \gamma_x \gamma_y$. Our approach is to select λ_1 , λ_2 and then find ϕ_x , ϕ_y to satisfy $\phi_x + \phi_y = \lambda_1 + \lambda_2$, $\phi_x \phi_y = \lambda_1 \lambda_2 + \gamma_x \gamma_y$. This requires solving a quadratic equation. If we denote the solutions as r_+ , r_- , then we get two cases: $(\phi_x, \phi_y) = (r_+, r_-)$ and $(\phi_y, \phi_x) = (r_+, r_-)$. This leaves us to select γ_x , γ_y .

In appendix B we show that $F_{Y \to X} = ln \frac{\sigma_x^2}{\sigma_a^2} \ge ln(1+\xi_x)$ where $\xi_x = \gamma_x^2(1-\rho^2)$. Similarly, $F_{X \to Y} \quad ln(1+\xi_y)$ where $\xi_y = \gamma_y^2(1-\rho^2)$. But we also show that $\xi_x = 0 \Rightarrow F_{Y \to X} = 0$ and ξ_y

 $= 0 \Rightarrow F_{X \to Y} = 0$, so we select ξ_x , ξ_y , thereby setting a lower bounds to $F_{Y \to X}$, $F_{X \to Y}$. This seems to be the best one can do, and as we see below, it works quite well. So given ξ_x , ξ_y ,

compute $\gamma_x = \pm \frac{\sqrt{\xi_x}}{\sqrt{1-\rho^2}}$ and $\gamma_y = \pm \frac{\sqrt{\xi_y}}{\sqrt{1-\rho^2}}$. This gives four cases and multiplied by the two cases above yields eight cases.

This is not quite the end of the story since the γ_x , γ_y values need to be consistent with the ϕ_x , ϕ_y values. Specifically the quadratic equation to be solved for ϕ_x , ϕ_y must have real roots. Thus, the discriminant must be 0. So $(\phi_x + \phi_y)^2 - 4(\phi_x \phi_y) = (\lambda_1 + \lambda_2)^2 - 4(\lambda_1 \lambda_2 + \gamma_x \gamma_y) = (\lambda_1 - \lambda_2)^2 - 4\gamma_x \gamma_y$ 0. There are four cases: two with real roots and two with complex roots.

If λ_1 , λ_2 are real, then we require

 $(\lambda_1 - \lambda_2)^2 \ge 4\gamma_x \gamma_y \Rightarrow (\lambda_1 - \lambda_2)^2 \ge 4sign (\gamma_x \gamma_y) \frac{\sqrt{\xi_x \xi_y}}{1 - \rho}$. This always holds if $sign(\gamma_x \gamma_y)$ 0. If $sign(\gamma_x \gamma_y) > 0$, we have a binding constraint that restricts the sizes of ξ_x , ξ_y .

If λ_1 , λ_2 are complex conjugates, then $(\lambda_1 - \lambda_2)^2$ is negative. If $sign(\gamma_x \gamma_y) = 0$, then the condition never holds. If $sign(\gamma_x \gamma_y) < 0$, then there is a binding constraint that restricts the sizes of ξ_x , ξ_y . In particular, note that if $sign(\gamma_x \gamma_y) = 0$, then one cannot have complex roots for *A*. We now use this design procedure to illustrate reverse causality.

8.2 Computation

We describe the steps used to generate the results below. We assume the state-space model

for $z_t = \begin{pmatrix} x_t \\ y_t \end{pmatrix}$ comes in ISS form. Since standard state-space subspace model-fitting algorithms (Larimore, 1983; van Overschee & de Moor, 1996; Bauer, 2005) generate ISS models, this is a reasonable assumption. Otherwise we use result 1 to generate the corresponding ISS model.

Given a sampling multiple *m*, we first use theorem 3 to generate the subsampled ISS model and, hence, $\sum_{\varepsilon}^{(m)}$. To obtain the GEMs, we use theorem 1 to generate the marginal models for $x_b y_t$ yielding $\Omega_x^{(m)}$, $\Omega_y^{(m)}$. And now $F_{Y \to X}^{(m)}$, $F_{X \to Y}^{(m)}$ are gotten from formulas 5.8 and 5.9 and the comment following theorem 5a.

8.3 Scenario Studies

We now illustrate the various results above with some bivariate simulations.

Example 1: GEMs Decline Gracefully—Here, for the underlying process, *y* pushes *x* much harder than *x* pushes *y*. This pattern is roughly preserved with slower sampling, but the relative strengths change as seen in Table 1.

Example 2: GEMs Reverse—In this case, the underlying processes push each other with roughly equal strength. But subsampling yields a false picture with *x* pushing *y* much harder than the reverse (see Table 2).

Example 3: Near Equal Strength Dynamics Becomes Nearly Unidirectional—In this case, the underlying relation is one of near equal strength feedback interconnection. But as Table 3 shows, almost immediately a very unequal relation appears under subsampling which soon decays to a near unidirectional relation.

Example 4: Near Unidirectional Dynamics Becomes Near Equal Strength—In

this case a near unidirectional dynamic relation immediately becomes one of significant but unequal strengths and then one of near equal strength (see Table 4).

There is nothing pathological about these examples, and using the design procedure developed above, it is easy to generate other similar kinds of examples. They make it emphatically clear that GC cannot be reliably discerned from downsampled data.

In independent work, Seth et al. (2013) have shown simulations where downsampling distorts GC. But their work is purely empirical, and they do not have the theoretical framework developed here to design specific types of distortion as we have done.

9 Application to fMRI

In this section we apply the above results to fMRI, but to do so, we need to describe fMRI data generation schemes (dgs's); we give three such schemes The following feature is common to all of them.

We suppose there is an underlying neural activity (NA) 2-node circuit operating on a fast timescale of order 50 ms and that it is described by a state-space model. Specifically, the (possibly vector) NA signals at each node are $x_b y_b$ which we collect together into a joint

NA vector signal $z_t = \begin{pmatrix} x_t \\ y_t \end{pmatrix}$, which is described by a state-space model such as equation 2.1 or 2.3.

• Filter (F). The BOLD signal recorded at a particular voxel is generated as follows. The NA z_t is filtered by (a spatially varying) HRF to generate a fast BOLD signal $\overline{z}_{b,t}$ —specifically

$$\overline{z}_{b,t} = \begin{pmatrix} \overline{x}_{b,t} \\ \overline{y}_{b,t} \end{pmatrix} = \begin{bmatrix} H_{xx}(L) & 0 \\ 0 & H_{yy}(L) \end{bmatrix} z_t$$

• Filter then downsample (F-DS). The BOLD signal recorded at a particular voxel is generated in two stages. The first stage is the same as F. In the second stage, the scanner samples this signal delivering a downsampled (or slow) BOLD signal $z_{b,k} = \overline{z}_{b,mt}$ —for example, if the NA timescale is 50 ms and the downsampled time scale is 200 ms, then m = 4.

• Downsample then filter (DS-F). Here we suppose the NA z_t is down-sampled to a slow timescale giving a slow NA $\overline{z}_k = z_{mt}$. This is then filtered by (a slow spatially varying) HRF to generate the slow BOLD signal $z_{h,k}$.

Note that F is the limiting case of F-DS where fMRI is recorded on the same timescale as NA. Clearly F-DS makes physiological sense, whereas DS-F does not. Indeed F-DS is advocated by Deshpande et al. (2010), who provide further physiological references to back it up. However, below we will see that DS-F seems to be behind some of the processing methods used in practice.

The aim now is to try to solve the reverse GC problem; to determine the GCS of fast NA circuit z_t from the GCS of either the fast BOLD circuit $\overline{z_i}$ or the slow BOLD circuit $z_{b,k}$. To do this, we see from theorem 6, that we need to investigate the phase properties of the HRF.

9.1 HRFs Are Nonminimum Phase

We denote an HRF by h_r , $r=1, 2, \cdots$ and its transfer function as

 $h(L) = \sum_{1}^{\infty} h_r L^r = \sum_{1}^{\infty} h_r z^{-r} = H(z)$. We assume H(z) is bounded input, bounded output stable (Kailath et al., 2000), for which the necessary and sufficient condition is

$$B = \sum_{1}^{\infty} |h_r| < \infty.$$

In Figure 1 we show plots of canonical HRFs (with default parameter choices) used in SPM and Glover (1999). They exhibit three main features found in practice:

- **1.** Refractory period: Characterized by a time to onset τ_o
- 2. Overshoot—characterized by a peak time τ_p
- 3. Undershoot

Also there are persistent reports in the literature of HRFs with an initial small dip immediately following the refractory period (e.g., Miezin, Maccotta, Ollinger, Petersen, & Buckner, 2000).

A number of results in the control engineering literature show how nonminimum phase (NMP) zeros of the transfer function imply overshoot and undershoot of an impulse response (see Damm & Muhirwa, 2014). But none of these works provide converse results, showing what impulse response shape implies about transfer function zeros.

We give a preliminary general result and then two converse results.

Theorem 8a—Any HRF with a refractory period is NMP.

Proof—We can write the transfer function as $h(L) = L^{\tau} o g(L)$ where g(L) is a causal transfer function. Even if g(L) is MP, then h(L) is NMP because $L^{-\tau} o$ is not a causal operator.

Certainly the time to onset can be estimated, and then the BOLD signal can be shifted in time to align the start time with the stimulus start time. This effectively corrects the HRF for

the refractory period. But this makes sense only if it is known that the refractory period is due to information transmission delay. If it is due to the GCS, then such alignment, being noncausal, will destroy the GCS.

To continue, we now assume the HRF has been corrected for the refractory period so that the corrected transfer function is $H_c(z) = z \sum_{1}^{\infty} h_r z^{-r}$.

Theorem 8b—Dip: if $h_1 < 0$ and H(1) > 0, then the HRF is NMP.

Note that

 $H_c(1)=H(1)=\sum_{1}^{\infty}h_r=-$ area above the dip+area under the overshoot – area above the undershoot . The fact that the overshoot area (far) exceeds the sum of the other two areas (so that H(1) > 0) seems to be a universal property of actual HRFs.

Proof—We prove that $H_c(z)$ has a real zero outside the UC. We can set z = R, which is real. Then consider that $H_c(R) = RH(R) = R \sum_{1}^{\infty} h_r R^{-r} = h_1 + \sum_{2}^{\infty} h_r R^{-(r-1)}$. For R = 1, the second term is bounded by $R^{-1} \sum_{2}^{\infty} |h_r| < R^{-1}B$ and $\rightarrow 0$ as $R \rightarrow \infty$. Since $h_1 < 0$, then for R large enough, RH(R) < 0 Since 1H(1) > 0, this means there must exist $R_o > 1$ with $R_oH(R_o) = 0$ as required.

Theorem 8c: No dip: if $h_1 > 0$ and H(-1) > 0, then the HRF is NMP.

Note that $H(-1) = \sum_{1}^{\infty} (-1)^r h_r = \sum_{m=1}^{\infty} (h_{2m} - h_{2m-1})$ and H(-1) > 0 is satisfied by the default canonical HRFs. Further, because HRFs are very smooth, the successive differences $h_{2m} - h_{2m-1}$ will be very small. H(-1) can be split into four pieces: a positive sum up to the peak time, a negative sum from there to the zero crossing, a positive sum to the trough of the undershoot, and a negative sum from there back up to the the time axis. We thus expect H(-1) to be small.

Proof—We prove $H_c(z)$ has a complex zero outside the UC. Put $z = Re^{j\pi} = -R$. Then consider that $-H_c(-R) = RH(-R) = R\sum_{1}^{\infty} h_r(-R)^{-r} = -h_1 + \sum_{2}^{\infty} h_r(-1)^r R^{-(r-1)}$. For R 1, the second term is bounded by $R^{-1}\sum_{2}^{\infty} |h_r| < R^{-1}B$, which $\rightarrow 0$ as $R \rightarrow \infty$ and so since $-h_1 < 0$, then for R large enough, RH(-R) < 0. Since 1H(-1) > 0, there must thus exist $R_o > 1$ for which $R_oH(-R_o) = 0$ and the result is established.

Since the condition H(-1) > 0 and more generally the possibility of complex zeros outside the UC is so crucial, more flexible modeling of the HRF is called for. Recent approaches include Khalidov, Fadili, Lazeyras, Ville, and Unser (2011), Zafer, Blu, and Ville (2014), Wu et al. (2013), Sreenivasan, Havlicek, and Deshpande (2015) and our approach using Laguerre-polynomials (Cassidy, Long, Rae, & Solo, 2012).

As noted in section 1, the current practice is to first apply a deconvolution to the BOLD signals before carrying out a GC analysis. As discussed in section 9.3, this makes sense only for DS-F.

We now see from the deconvolution example 6 C that even if Wiener deconvolution is applied to noise-free data $z_{b,k}$, it will distort the GCS because it necessarily involves noncausal filtering. This is discussed further in section 9.3.

9.3 Reverse GC Problem

We discuss approaches under F, F-DS, and DS-F.

F—There is only one approach possible here. If the HRF is minimum phase, then the fast filtering will not affect the GCS and the first stage would consist of a GC analysis of the fast BOLD data. That is, fit an SS model to $\overline{z}_{b,t}$; find submodels for $\overline{x}_{b,t}$, $\overline{y}_{b,t}$ using DARE (see theorem 1); and compute the GEMS $F_{Y \to X}$, F_{YX} , $F_{X \to Y}$ using the expressions in theorem 5a. This would deliver the GCS of the fast BOLD circuit and, hence, of the NA circuit. But in practice, the BOLD data will be noisy, and an attempt to reduce the impact of the noise by deconvolution will distort the GCS. If the HRF is nonminimum phase, then this GCS would be distorted and nothing can be done to overcome that.

F-DS—Again there is only one approach possible here. If the HRF is minimum phase, then the fast filtering will not affect the GCS and the first stage would consist of a GC analysis of the slow BOLD data—that is, fit an SS model to $z_{b,k}$; find submodels for $x_{b,k}$, $y_{b,k}$ using DARE (see theorem 1); and compute the GEMS $F_{Y \to X}$, $F_{Y,X}$, $F_{X \to Y}$ using the expressions in theorem 5a. This would deliver the GCS of the slow BOLD circuit. The second stage is to try to get the GCS of the fast BOLD circuit. But now we see that it must fail since the results of the reverse GC study in section 8 demonstrate that we cannot use the GCS of the slow BOLD circuit to discern the GCS of the fast BOLD circuit. So we cannot get at the GCS of the fast NA circuit.

Notice that in F-DS there is no role for deconvolution.

If the HRF, is NMP then the GCS will be distorted and there is nothing we can do to overcome this. If we apply deconvolution, then the analysis above shows it must fail since the Wiener filter is NMP.

DS-F—There are two approaches here. The direct approach has two stages. If the HRF is minimum phase, then the slow filtering will not affect the GCS and the first stage would consist of a GC analysis of the slow BOLD data—that is, fit an SS model to $z_{b,k}$; find submodels for $x_{b,k}$, $y_{b,k}$ using DARE (see theorem 1); and compute the GEMS $F_{Y \to X}$, $F_{Y,X}$, $F_{X \to Y}$ using the expressions in theorem 5a. This would deliver the GCS of the slow NA circuit since by theorem 6, the GEMS would be the same as for the slow BOLD data. The second stage is to try to get the GCS of the fast NA circuit. But now we see that it must fail, since the results of the reverse GC study in section 8 demonstrate that we cannot use the GCS of the slow NA circuit to discern the GCS of the fast NA circuit.

The indirect approach has three stages. If the slow HRF is minimum phase, then in the zeroth stage, we could inverse filter (deconvolve) the slow BOLD data to obtain the slow NA. In the first stage, we obtain its GCS using the procedure described in the first stage of the direct approach. But for the second stage, we are in the same situation as in the direct approach and must fail to discern the GCS of the fast NA.

In practice, slow timescale deconvolution has gained favor recently. It is done using a Wiener filter (Glover, 1999) since this deals with possible noise (which we have ignored in our discussion). But as discussed above, this must fail since the Wiener filter is NMP.

We see here that only in DS-F is there a role for slow timescale deconvolution.

In sum we conclude that under any data-generating scheme, GC analysis must fail.

We note that there is an extensive discussion of GC in fMRI in Rodrigues and Andrade (2015). Unfortunately this work relies on the erroneous method of Chen et al. (2006) for computing GEMs (see the remarks following theorem 5b), and so its results are flawed.

10 Conclusion

This letter has given a theoretical and computational analysis of the use of Granger causality (GC) and applied the results to fMRI. There were two main issues: the effect of downsampling and the effect of hemodynamic filtering on Granger causal structure (GCS). To deal with these issues, a number of novel results in multivariate time series and Granger causality were developed via state-space methods as follows:

- 1. Computations of submodels via the DARE (see theorems 1 and 4)
- 2. Computation of downsampled models via the DARE (see theorem 3)
- 3. Reliable computation of GEMs via the DARE (see theorems 5a and 5b)
- 4. Effect of filtering on GEMs (see theorem 6)

Theorem 6 showed that nonminimum phase filters distort GCS. Using the results in the first three methods, we were able to develop, in section 8, a framework for generating downsampling-induced spurious Granger causality on demand and provided a number of illustrations.

In section 9 these results were applied to fMRI. To do this we described three data generating processes for fMRI:

- 1. F where fMRI time series are filtered neural activity (NA)
- 2. F-DS where NA is filtered, then downsampled to give fMRI time series
- 3. DS-F where NA is downsampled, then fitered to give fMRI time series

While F-DS has strong physiological justification, DS-F seems to be behind a number of approaches in the literature. We showed that HRFs are non-minimum phase due to a refractory period. If that period can be corrected for, the phase character of the HRF has to

be checked on a case-by-case basis. Certainly default canonical HRFs are nonminimum phase, and so more flexible HRF modeling will be needed.

Then we discussed the reverse GC problem under each data-generating process. In each case, we concluded that GC analysis could not work.

All this leads to the conclusion that linear Granger causality analysis of fMRI data cannot be used to discern neuronal level–driving relationships. Not only is the timescale too slow, but even with faster sampling, the non-minimum phase aspect of the HRF and the noise will still compromise the method.

Future work would naturally include an extension of the Granger causality results to handle the presence of a third vector time series and extensions to deal with time-varying Granger causality. Nongaussian versions could mitigate the nonminimum phase problem to some extent, but there does not seem to be any evidence for the nongaussianity of fMRI data. Extensions to nonlinear Granger causality are currently of great interest but need considerable development.

Acknowledgments

In the May 2016 issue of *Neural Computation* (Volume 28, Number 5, pp. 914–949), the letter "State-Space Analysis of Granger-Geweke Causality Measures with Application to fMRI" by Victor Solo was printed with incomplete acknowledgments.

The published acknowledgment reads:

This work was partially supported by NIH grant P41EB015896.

The complete acknowledgment should read:

This work was partially supported by the NIH grant P41EB015896. The author would first like to thank Martinos Center colleagues Dr. Bruce Rosen and Dr. Robert Weiskoff who raised crucial questions during a joint discussion of Granger causality and fMRI (held in 1998), which ultimately led to this paper. Thanks are also due to Dr. Graham Goodwin who asked a critical question during an IEEE Control Engineering conference presentation that led to the work on the reverse problem. Special thanks are due to a referee who posed a series of critical questions that led to a major expansion of section 9 into its current form. Finally we acknowledge discussions with Martinos Center colleagues Dr. Mark Vangel, Dr. Fa-Hsuan Lin, and Dr. Matti Håmålåinen.

Appendix A: Stabilizability, Detectability, and DARE

In this appendix we restate and modify for our purposes some standard state-space results. We rely mostly on Kailath et al. (2000, appendixes E and C).

We denote an eigenvalue of a matrix by λ and a corresponding eigenvector by q. We say λ is a stable eigenvalue if $|\lambda| < 1$; otherwise, λ is an unstable eigenvalue.

A.1 Stabilizability

The pair (A, B) is controllable if there exists a matrix G so that A-BG is stable (i.e., all eigenvalues of A-BG are stable). (A, B) is controllable iff any of the following conditons hold,

• Controllability matrix: $C = [B, AB, ..., A^{n-1}B]$ has rank *n*.

- Rank test: $rank[\lambda I A, B] = n$ for all eigenvalues λ of A.
- PBH test: There is no left eigenvector of A that is orthogonal to B, that is, if $q^T A = \lambda q^T$, then $q^T B = 0$.

The pair (*A*, *B*) is stabilizable if $rank[\lambda I - A, B] = n$ for all unstable eigen-values of *A*. Three useful tests for stabilizability are:

- PBH test: (A, B) is stabilizable iff there is no left eigenvector of A corresponding to an unstable eigenvalue that is orthogonal to B, that is, if $q^T A = \lambda q^T$ and $|\lambda| 1$, then $q^T B = 0$.
- (A, B) is stabilizable if (A, B) is controllable.
- (*A*, *B*) is stabilizable if *A* is stable.

A.2 Theorem DARE

See Kailath et al. (2000, theorem E6.1, lemma 14.2.1, section 14.7).

Under conditions N,St,De, the DARE has a unique positive semidefinite solution P, which is stabilizing (i.e., $A_s - K_sC$ is a stable matrix). Further if we initialize $P_0 = 0$, then P_t is nondecreasing and $P_t \rightarrow P$ as $t \rightarrow \infty$.

Remarks

- i. " $A_s K_s C$ stable" means that A KC is stable (see below), and this implies that (A, K) is controllable (see below).
- ii. Since V R, then $N \Rightarrow V$ is positive definite.

Proof of Remark i—We first note (taking limits in) (Kailath et al., 2000, equation 9.5.12) $K_s = K - SR^{-1}$. We then have $A_s - K_sC = A - SR^{-1}C - (K - SR^{-1})C = A - KC$. So $A_s - K_sC$ is stable iff A - KC is stable. But then (A, K) is controllable.

A.3 Detectability

The pair (A, C) is detectable if (A^T, C^T) is stabilizable.

Remarks

- i. If A_s is stable (all eigenvalues have modulus < 1), then St De automatically hold.
- **ii.** Condition De can be replaced with the detectability of (A, C) which is the way Kailath et al. (2000) state the result. We show equivalence below (this is also noted in a footnote in Kailath et al., 2000, section 14.7).

Proof of Remark ii

Suppose (A_{s}, C) is detectable but (A, C) is not. Then by the PBH test, there is a right eigenvector p of A corresponding to an unstable eigenvalue of A with $Ap = \lambda p, Cp = 0$. But then $A_sp = (A - SR^{-1}C)p = Ap = \lambda p$ while Cp = 0 which contradicts the detectbility of (A_s, C) . The reverse argument is much the same.

Proof of Result 1

Part a follows from the discussion leading to theorem DARE. Part b follows from the remarks after theorem DARE.

Appendix B: GEMs for Bivariate VAR(1)

Applying formula 5.5 and reading off H_{XX} from the VAR(1) model yields

$$\begin{split} f_{X}(\lambda) &= [(1+\phi_{y}^{2}-2\phi_{y}cos(\lambda))\sigma_{a}^{2}+\gamma_{x}^{2}\sigma_{b}^{2}\\ &-2\rho\sigma_{a}\sigma_{b}(\gamma_{x}\phi_{y}-\gamma_{x}cos(\lambda))]/|D(e^{j\lambda})|^{2}\\ D(L) &= (1-\phi_{x}L)(1-\phi_{y}L)-\gamma_{x}\gamma_{y}L^{2}. \end{split}$$

This can clearly be written as an ARMA(2,1) spectrum $\sigma_x^2 |1 - \theta_x e^{-j\lambda}|^2 / |D(e^{j\lambda})|^2$ Equating coefficients gives

$$\begin{split} \gamma_0 = & \sigma_x^2 (1 + \theta_x^2) = (1 + \phi_y^2) \sigma_a^2 + \sigma_b^2 \gamma_x^2 - 2\rho \sigma_a \sigma_b \gamma_x \phi_y = \sigma_a^2 (1 + \xi_x + d_x^2), \\ & \gamma_1 = \sigma_x^2 \theta_x = \sigma_a^2 d_x, \end{split}$$

where $\xi_x = (1 - \rho^2)\gamma_x^2 \sigma_b^2$ and $d_x = \phi_y = \rho \gamma_x \frac{\sigma_b}{\sigma_a}$. We thus have $\theta_x^2 = \gamma_1^2 / \sigma_x^4$, and using this in the first equation gives, $\gamma_0 = \sigma_x^2 + \gamma_1^2 / \sigma_x^2$ or $\sigma_x^4 - \sigma_x^2 \gamma_0 + \gamma_1^2 = 0$. This has, of course, two solutions:

$$\sigma_x^2 = \frac{1}{2} \left(\gamma_0 \pm \sqrt{\gamma_0^2 - 4\gamma_1^2} \right), \\ \Rightarrow \frac{\sigma_x^2}{\sigma_a^2} = \frac{1}{2} \left(\frac{\gamma_0}{\sigma_a^2} \pm \sqrt{\frac{\gamma_0^2}{\sigma_a^4} - 4\frac{\gamma_1^2}{\sigma_a^4}} \right) \\ = \frac{1}{2} \left(1 + \xi_x + d_x^2 \pm \sqrt{(1 + \xi_x + d_x^2)^2 - 4d_x^2} \right)$$

Note that if $\xi_x = 0$, this delivers $\frac{\sigma_x^2}{\sigma_a^2} = \frac{1}{2}(1+d_x^2+\sqrt{(1-d_x^2)^2}=1)$.

We must choose the solution that ensures

$$|\theta_x| < 1 \equiv \frac{\gamma_1^2}{\sigma_x^4} < 1 \equiv \frac{\gamma_0}{\sigma_x^4} < 2 \equiv \gamma_0/\sigma_a^2 < 2\sigma_x^2/\sigma_a^2 \equiv 1 + \xi_x + d_x^2 < 1 + \xi_x + d_x^2 \pm \sqrt{(1 + \xi_x + d_x^2)^2 - 4d_x^2}$$

And so we must choose the + solution. Continuing, we now claim.

And so we must choose the + solution. Continuing, we now claim,

$$\begin{aligned} &\frac{\sigma_x^2}{\sigma_a^2} \ge \frac{1}{2} \left(1 + \xi_x + d_x^2 + \sqrt{\left(1 + \xi_x - d_x^2\right)^2} \right) \\ &= \frac{1}{2} \left(1 + \xi_x + d_x^2 + 1 + \xi_x - d_x^2 \right) = 1 + \xi_x. \end{aligned}$$

This follows if

 $(1+\xi_x+d_x^2)^2-4d_x^2 \ge (1+\xi_x-d_x^2)^2 \equiv (1+\xi_x+d_x^2)-(1+\xi_x-d_x^2)^2 \ge 4d_x^2 \equiv 2(1+\xi_x)2d_x^2 \ge 4d_x^2 \equiv 1+\xi_x \ge 1$, which holds.

Appendix C: Spectral Factorization

Suppose e_t is a white noise sequence with $E(e_t) = 0$, $var(e_t) = \Sigma$. Let G(L) be a stable causal, possibly nonminimum phase filter. Then $\overline{z}_t = G(L)\varepsilon_t$ has spectrum $f_{\overline{z}}(\lambda) = G(L)VG^T(L^{-1})$ where $L = exp(-j\lambda)$. We can then find a unique causal, stable minimum phase spectral factorization $f_{\overline{z}}(\lambda) = G_o(L)V_oG_o^T(L^{-1})$. Let V, V_o have Cholesky factorizations $V = JJ^T$, $V_o = J_oJ_o^T$ and set $G_c(L) = G(L)J$, $G_{o,c}(L) = G_o(L)J_o$. Then $f_{\overline{z}}(\lambda) = G_c(L)G_c^T(L^{-1}) = G_{o,c}(L)G_{o,c}^T(L^{-1})$. Since $G_{o,c}(L)$ is minimum phase, we can introduce the causal filter $E(L) = G_{o,c}^{-1}(L)G_c(L) \Rightarrow E(L)E^T(L^{-1}) = I$. Such a filter is called an *all pass* filter (Hannan & Deistler, 1988; Green, 1988). Now, $G_c(L) = G_{o,c}(L)E(L)$ or G(L) $= G_o(L)J_oE(L)J^{-1}$ (i.e., a decomposition of a nonminimum phase (matrix) filter into a product of a minimum phase filter and an all-pass filter. We can also write this as

 $G_o(L) = G(L)JE^{-1}(L)J_o^{-1} = G(L)JE^T(L^{-1})J_o^{-1}$ showing how the nonminimum phase filter is transformed to yield a spectral factor.

References

- Baccala L, Sameshima K. Partial directed coherence: A new concept in neural structure determination. Biol Cybernetics. 2001; 84:463–474.
- Bauer D. Estimating linear dynamical systems using subspace. Econometric Theory. 2005; 21:181–211.
- Bernasconi C, Konig P. On the directionality of cortical interactions studied by structural analysis of electrophysiological recordings. Biol Cyb. 1999; 81:199–210.
- Caines P. Weak and strong feedback free processes. IEEE Trans Autom Contr. 1976; 21:737-739.
- Caines P, Chan W. Feedback between stationary stochastic processes. IEEE Trans Autom Contr. 1975; 20:498–508.
- Cassidy B, Long C, Rae C, Solo V. Identifying fMRI model violations with Lagrange multiplier tests. IEEE Trans Medical Imaging. 2012; 31:1481–1492. [PubMed: 22542665]
- Chen Y, Bressler S, Ding M. Frequency domain decomposition of conditional Granger causality and application to multivariate neural field potential data. J Neuroscience Methods. 2006; 150:228–237.
- Christiano L, Eichenbaum M. Temporal aggregation and structural inference in macroeconomics. 1987Technical Report, Federal Reserve Bank of Minnesota
- Damm T, Muhirwa LN. Zero crossings, overshoot and initial undershoot in the step and impulse responses of linear systems. IEEE Trans Autom Contr. 2014; 59:1925–1929.
- David O, Guillemain I, Saillet S, Reyt S, Deransart C, Segebarth C, Depaulis A. Identifying neural drivers with functional MRI: An electrophysiological validation. PLOS Biology. 2008; 6:2683– 2697. [PubMed: 19108604]
- Deistler M, Peternell K, Scherer W. Consistency and relative efficiency of subspace methods. Automatica. 1995; 31:1865–1875.
- Deshpande G, Sathian K, Hu X. Effect of hemodynamic variability on Granger causality analysis of fMRI. NeuroImage. 2010; 52:884–896. [PubMed: 20004248]
- Ding M, Bressler S, Yang W, Liang H. Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: Data preprocessing, model validation, and variability assessment. Biol Cyb. 2000; 83:35–45.
- Dufour J, Renault E. Short and long run causality in time series: Theory. Econometrica. 1998; 66:1099–1125.
- Dufour J, Taamouti A. Short and long run causality measures: Theory and inference. Journal of Econometrics. 2010; 154:42–58.

- Edin F. Scaling errors in measures of brain activity cause erroneous estimates of effective connectivity. NeuroImage. 2010; 49:621–630. [PubMed: 19607927]
- Feinberg D, Yacoub E. The rapid development of high speed, resolution and precision in fMRI. NeuroImage. 2012; 62:720–725. [PubMed: 22281677]
- Geweke J. Temporal aggregation in the multiple regression model. Econo- metrica. 1978; 46:643-661.
- Geweke J. The measurement of linear dependence and feedback between multiple time series. J Amer Stat Assoc. 1982; 77:304–313.
- Geweke J. Measures conditional of linear dependence and feedback between time series. J Amer Stat Assoc. 1984; 79:907–915.
- Glover G. Deconvolution of impulse response in event-related bold fMRI. NeuroImage. 1999; 9:416–429. [PubMed: 10191170]
- Goebel R. Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. Magn Reson Imaging. 2003; 21:1251– 1261. [PubMed: 14725933]

Granger C. Economic processes involving feedback. Information and Control. 1963; 6:28-48.

- Granger C. Investigating causal relations by econometric models and cross- spectral methods. Econometrica. 1969; 37:424–438.
- Green M. All-pass matrices, the positive real lemma and unit canonical correlations between future and past. J Multiv Anal. 1988; 24:143–154.
- Hamilton, J. Time series analysis. Princeton, NJ: Princeton University Press; 1994.
- Handwerker D, Gonzalez-Castillo J, D'Esposito M, Bandettini P. The continuing challenge of understanding and modeling hemodynamic variation in fMRI. NeuroImage. 2012; 62:1017–1023. [PubMed: 22366081]
- Hannan, E., Deistler, M. Statistical theory of linear systems. New York: Wiley; 1988.
- Hannan E, Poskitt D. Unit canonical correlations between future and past. Ann Stat. 1988; 16:784–790.
- Havlicek M, Friston K, Jan J, Brazdil M, Calhoun VD. Dynamic modeling of neuronal responses in fMRI using cubature Kalman filtering. NeuroImage. 2011; 56(4):2109–2128. [PubMed: 21396454]
- Hsiao, C. Time series modeling and causal ordering of Canadian money, income and interest rates. In: Anderson, OD., editor. Time series analysis: Theory and practice. Amsterdam: North-Holland; 1982. p. 671-699.
- Kailath, T., Sayeed, A., Hassibi, B. Linear estimation. Upper Saddle River, NJ: Prentice Hall; 2000.
- Kaminski M, Blinowska K. A new method of the description of the information flow in the brain structures. Biol Cyb. 1991; 65:203–210.
- Kaufmann K, Stern D. Evidence for human influence on climate from hemispheric temperature relations. Nature. 1997; 388:39–44.
- Khalidov I, Fadili J, Lazeyras F, Ville D, Unser M. Activelets: Wavelets for sparse representation of hemodynamic responses. Signal Processing. 2011; 91:2810–2821.
- Kshirsagar, A. Multivariate analysis. New York: Marcel Dekker; 1972.
- Larimore, W. Proc American Control Conference. Piscataway, NJ: IEEE; 1983. System identification, reduced order filters and modeling via canonical variate analysis; p. 445-451.
- Lutkepohl, H. Introduction to multivariate time series analysis. Berlin: Springer-Verlag; 1993.
- Magnus, J., Neudecker, H. Matrix differential calculus with applications in statistics and econometrics. New York: Wiley; 1999.
- Marcet, A. Temporal aggregation of economic time series. In: Hanson, LP., Sargent, TJ., editors. Rational expectations matrix. Boulder, CO: Westview Press; 1991. p. 237-281.
- Middleton, R., Goodwin, G. Digital control and estimation: A unified approach. Upper Saddle River, NJ: Prentice Hall; 1990.
- Miezin FM, Maccotta L, Ollinger JM, Petersen SE, Buckner RL. Characterizing the hemodynamic response: Effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. NeuroImage. 2000; 11:735–759. [PubMed: 10860799]

Solo

Pandit, S., Wu, S. Time series and systems analysis with applications. New York: Wiley; 1983.

- Pascual-Marqui, R., Biscay, R., Bosch-Bayard, J., Lehmann, D., Kochi, K., Yamada, N., Sadato, N. Isolated effective coherence (ICOH): Causal information flow excluding indirect paths. Zurich: KEY Institute for Brain-Mind Research, University of Zurich; 2014. Technical ReportarXiv: 1402.4887
- Pierce D, Haugh L. Causality in temporal systems characterizations and a survey. Journal of Econometrics. 1977; 5:265–293.
- Rissannen J, Wax M. Measures of mutual and causal dependence between two time series. IEEE Trans Inf Thy. 1987; 33:598–601.
- Rodrigues J, Andrade A. Lag-based effective connectivity applied to fMRI: A simulation study highlighting dependence on experimental parameters and formulation. NeuroImage. 2015; 89:358–377.
- Roebroeck A, Formisano E, Goebel R. Mapping directed influence over the brain using Granger causality. NeuroImage. 2005; 25:230–242. [PubMed: 15734358]
- Roebroeck A, Formisano E, Goebel R. The identification of interacting networks in the brain using fMRI: Model selection, causality and deconvolution. NeuroImage. 2011; 58:296–302. [PubMed: 19786106]
- Seber, G. Multivariate observations. New York: Wiley; 1984.
- Seth A, Chorley P, Barnett L. Granger causality analysis of fMRI bold signals is invariant to hemodynamic convolution but not downsampling. NeuroImage. 2013; 65:540–555. [PubMed: 23036449]
- Sims C. Discrete approximations to continuous time distributed lags in econometrics. Econometrica. 1971; 39:545–563.
- Sims C. Money, income and causality. Am Econ Rev. 1972; 62:540-552.
- Solo, V. Topics in advanced time series analysis. In: del Pino, G., Rebolledo, R., editors. Lectures in probability and statistics. Berlin: Springer-Verlag; 1986. p. 165-328.
- Sreenivasan K, Havlicek M, Deshpande G. Nonparametric hemodynamic deconvolution of fMRI using homomorphic filtering. IEEE Trans Medical Imaging. 2015; 34:1155–1163. [PubMed: 25531878]
- Suna L, Wang M. Global warming and global dioxide emission: An empirical study. J Environmental Management. 1996; 46(4):327–343.
- Takahashi D, Baccala L, Sameshima K. Partial directed coherence asymptotics for VAR processes of infinite order. International Journal of Bioelectromagnetism. 2008; 10:31–36.
- Telser L. Discrete samples and moving sums in stationary stochastic processes. J Amer Stat Assoc. 1967; 62:484–499.
- Triacca U. Is Granger causality analysis appropriate to investigate the relationship between atmospheric concentration of carbon dioxide and global surface air temperature? Theoretical and Applied Climatology. 2005; 81:133–135.
- Valdes-Sosa P. Spatio-temporal autoregressive models defined over brain manifolds. Neuroinformatics. 2004; 2:239–250. [PubMed: 15319519]
- Valdes-Sosa P, Roebroeck A, Daunizeau J, Friston K. Effective connectivity: Influence, causality and biophysical modeling. NeuroImage. 2011; 58:339–361. [PubMed: 21477655]
- van Overschee, P., de Moor, B. Subspace identification for linear systems. Dordrecht: Kluwer; 1996.
- Wu G, Liao W, Stramaglia S, Ding J, Chen H, Marinazzo D. A blind deconvolution approach to recover effective connectivity brain networks from resting state fMRI data. Medical Image Analysis. 2013; 17:365–374. [PubMed: 23422254]
- Yamashita O, Sadato N, Okada T, Ozaki T. Evaluating frequency-wise directed connectivity of bold signals applying relative power contribution with the linear multivariate time-series models. NeuroImage. 2005; 25:478–490. [PubMed: 15784427]
- Zafer, D., Blu, T., Ville, D. Proc IEEE ISBI. Piscataway, NJ: IEEE; 2014. Detecting spontaneous brain activity in functional magnetic resonance imaging using finite rate of innovation; p. 1047-1050.



Figure 1. Canonical HRFs.

Author Manuscript

GEMs for Various Sampling Intervals for
$$(\lambda_1, \lambda_2, \xi_x, \xi_y, \rho) = (-.95 \exp(j \times .1), -.95 \exp(-j \times .1), 1.5, .2, .2) \Rightarrow A = \begin{pmatrix} -.204 & -1.24 \\ .452 & -1.69 \end{pmatrix}$$
.

u	1	2	3	4	5	9	10	20	30	40
$T_{Y \to X}$	1.38	1.66	1.41	1.17	0.99	0.86	0.551	0.15	0.00	0.01
$T_{X \to Y}$	0.20	0.25	0.29	0.31	0.32	0.32	0.29	0.11	0.00	0.01

<

m	1	2	3	4	5	9	10	20	30	40
$F_{Y \to X}$	0.93	0.88	0.77	0.683	0.62	0.57	0.42	0.13	0.00	0.01
$F_{X \to Y}$	1.05	1.82	2.01	1.80	1.53	1.30	0.75	0.18	0.00	0.02

Table 3

 $\begin{pmatrix} -1.45 & -1.18 \\ -.84 & -1.16 \end{pmatrix}$ A=GEMs for Various Sampling Intervals for $(\lambda_1, \lambda_2, \xi_x, \xi_y, \rho) = (.995, -.7, 1, .5, .7) \Rightarrow$

m	1	2	3	4	5	9	10	20	30	40
$F_{Y \to X}$	1.49	0.05	0.25	0.06	0.11	0.05	0.04	0.02	0.01	0.01
$F_{X \to Y}$	1.69	0.17	0.64	0.26	0.47	0.29	0.29	0.21	0.16	0.13

Author Manuscript

Author Manuscript

Table 4

GEMs for Various Sampling Intervals for
$$(\lambda_1, \lambda_2, \xi_x, \xi_y, \rho) = (.99 \exp(j \times .25), .95 \exp(-j \times .25), .1, 3, -.8) \Rightarrow A = \begin{pmatrix} 1.883 & -0.408 \\ 2.236 & 0.036 \end{pmatrix}$$

	7	33	4	S	9	10	20	30	40
).28	C	.38	0.43	0.45	0.46	0.31	0.45	0.41	0.17
2.38	, 	.62	1.24	1.02	0.86	0.37	0.53	0.45	0.18

/