Journal of Neuroscience Methods 201 (2011) 404–419

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

Behaviour of Granger causality under filtering: Theoretical invariance and practical application

Lionel Barnett*, Anil K. Seth

Sackler Centre for Consciousness Science and School of Informatics, University of Sussex, Brighton BN1 9QJ, UK

ARTICLE INFO

Article history:
Received 2 June 2011
Received in revised form 1 August 2011
Accepted 4 August 2011

PACS:
87.19.L–
87.19.lj
87.10.Mn

Keywords:
Granger causality
Digital filtering
Vector autoregressive modelling
Time series analysis

ABSTRACT

Granger causality (G-causality) is increasingly employed as a method for identifying directed functional connectivity in neural time series data. However, little attention has been paid to the influence of common preprocessing methods such as filtering on G-causality inference. Filtering is often used to remove artifacts from data and/or to isolate frequency bands of interest. Here, we show [following Geweke (1982)] that G-causality for a stationary vector autoregressive (VAR) process is fully invariant under the application of an arbitrary invertible filter; therefore filtering cannot and does not isolate frequency-specific G-causal inferences. We describe and illustrate a simple alternative: integration of frequency domain (spectral) G-causality over the appropriate frequencies ("band limited G-causality"). We then show, using an analytically solvable minimal model, that in practice G-causality inferences often do change after filtering, as a consequence of large increases in empirical model order induced by filtering. Finally, we demonstrate a valid application of filtering in removing a nonstationary ("line noise") component from data. In summary, when applied carefully, filtering can be a useful preprocessing step for removing artifacts and for furnishing or improving stationarity; however filtering is inappropriate for isolating causal influences within specific frequency bands.

1. Introduction

A key theme in contemporary neuroscience is to move from localisation of function to characterisation of functional networks. In particular, analysis methods aimed at extracting directed functional (i.e., causal) connectivity from neural signals are increasingly in demand.[1] G-causality analysis is widely employed to identify causal connectivity in neural time series data. G-causality is a statistical measure of causality based on precedence and predictability. Put simply, if a variable A contains information that helps predict another variable B, better than can be done knowing only the past of B itself, then A is said to "G-cause" B. The concept has typically been operationalised in the context of linear VAR models and its uptake within neuroscience has been facilitated by the appearance of dedicated software toolboxes implementing the methods (Seth, 2010; Cui et al., 2008). However, the interaction of G-causality with standard data preprocessing procedures is not well understood and presents a possibly serious confound to many applications. In this paper, we focus on the effects of (temporal) filtering on G-causality. This is a crucial issue since filtering is often applied semi-automatically as a preprocessing step in many analyses. Most applications of filtering attempt to achieve one (or both) of two objectives: (i) removal of artifacts such as electrical line noise and (non-neural) physiological influences, and (ii) isolation of effects within a specific frequency band [e.g., the beta or gamma ranges in M/EEG (Pollonini et al., 2010; Wilson and Yan, 2010)]. Anticipating our results, we show that G-causality is theoretically invariant under the application of arbitrary (invertible) multivariate filters, and so cannot achieve the second objective. However, the invariance holds strictly for *stationary* data—stationarity being a prerequisite for G-causality analysis—so that filtering can be useful for artifact removal if it is able to render a previously nonstationary time series stationary. In practice, filtering can pose challenges for the effective estimation of the autoregressive models on which G-causality is based, hence the need for its careful application in the context of achieving or improving stationarity. Although our analysis is targeted at "explicit" filtering imposed by an experimenter as

Abbreviations: G-causality, Granger causality; iid, identically and independently distributed; MVGC, multivariate Granger causality; VAR, vector autoregressive; VMA, vector moving average; VARMA, vector autoregressive moving average; FIR, finite impulse response; IIR, infinite impulse response; OLS, ordinary least squares; AIC, Akaike information criterion; BIC, Bayesian information criterion; CV, cross-validation; EEG, electroencephalography; MEG, magnetoencephalography; fMRI, functional magnetic resonance imaging; BOLD, blood oxygen level dependent; HRF, hemodynamic response function; DTF, directed transfer function; PDC, partially directed coherence.

* Corresponding author. Tel.: +44 1273 699246.
E-mail addresses: l.c.barnett@sussex.ac.uk (L. Barnett), a.k.seth@sussex.ac.uk (A.K. Seth).

[1] We prefer the term *causal connectivity*, a description of the data, to *effective connectivity*, which implies a model of the underlying mechanism; see Bressler and Seth (2011).

0165-0270/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jneumeth.2011.08.010

a data (pre)processing stage, our results may also have implications for "implicit" filtering that may arise as a result of physiological processes intervening between neural variables and observables, for example as manifest in the hemodynamic BOLD signal measured using fMRI.

In his seminal 1982 paper Geweke (1982) noted, but did not justify or explore, the invariance of G-causality under filtering via the somewhat oblique aside "[G-causality] is invariant with respect to scaling of X and Y; in fact it remains unchanged if X and Y are pre-multiplied by different invertible lag operators." Perhaps because there is no explicit reference to "filtering" this note appears to have been overlooked as G-causality has been taken up within neuroscience. More recently, researchers have worried that filtering does in fact affect G-causality (Florin et al., 2010; Seth, 2010). A recent study by Florin et al. (2010) suggested that application of filtering to neural data disturbs the information content and time ordering of the data, leading to spurious and missed causalities (Type I and Type II errors, respectively). Their conclusion is based on the correct observation that filtering in general alters the regression coefficients of VAR models of the data. They then show using numerical simulations that filtering induces Type I and Type II errors in sample.[2] However, they did not make any analytical connection between the two observations. In fact, as we argue, the errors observed in simulation by Florin et al. derive from the difficulties inherent in fitting VAR models to filtered data, not from the filtering process *per se*. In particular, filtering generally induces a large increase in the empirical model order (the number of lagged observations incorporated into a VAR), leading to model mis-specification given limited data.

Our paper is organised as follows: in Section 2 we define G-causality in both the time and frequency domains, for unconditional and conditional situations, and for both univariate and multivariate (block, ensemble) variables. We also discuss estimation for finite-sample empirical data and significance testing. Readers familiar with the mathematical basis of G-causality may wish to skip this section, referring to it where needed for notation. In Section 3 we demonstrate analytically the invariance of G-causality under the application of an (almost) arbitrary stable, invertible, multivariate filter. The invariance is completely general, applying to all the varieties of G-causality just mentioned. We then consider issues arising in empirical estimation of G-causality, suggesting several reasons why filtering may corrupt empirical estimates despite the theoretical invariance. As mentioned, these turn principally on an increase in empirical model order induced by filtering; filtering may also induce near-nonstationarity and other numerical instabilities. Consequently, we argue that (i) filtering can be useful for preprocessing nonstationary (or near-nonstationary) time series and (ii) estimation of G-causality within specific frequency bands can be accomplished by integrating the frequency domain G-causality over the appropriate frequencies ("band limited G-causality"). Section 4 introduces a minimal VAR system for which G-causalities can be obtained analytically. We use this model to test how empirical estimates of G-causality are influenced by both FIR and IIR filters. We compare estimates of model order for unfiltered and filtered processes, showing a large increase in optimal (empirical) model order following filtering, as well as an increase in the likelihood of unstable VAR models. We then analyse the effects of model order and filtering on statistical significance testing, showing [consistent with Florin et al. (2010); Seth (2010)] increases in both Type I and Type II errors after filtering. We explain this result by showing a strong association between increased error rates and an increase in VAR model order entailed by filtering. Based on these findings, we demonstrate a useful example of filtering to remove line noise. Finally, we show that band-limited G-causality on unfiltered data correctly identifies frequency specific causal interactions, whereas G-causality on filtered data does not. Our conclusions are summarised and discussed in Section 5.

[2] A similar corruption of G-causality inferences by filtering was shown in another set of recent simulations (Seth, 2010).

2. Multivariate G-causality (MVGC)

Consider a covariance-stationary, n variable, VAR(p) process $\boldsymbol{U}_t$ (the "universe" of measurable variables) specified by the model[3]

$$\sum_{k=0}^{p} A_k \cdot \boldsymbol{U}_{t-k} = \boldsymbol{\varepsilon}_t \tag{1}$$

for $-\infty < t < \infty$, where the $n \times n$ square matrices A_k, $k = 0, 1, 2, \ldots, p$ are the regression coefficients with $A_0 \equiv I$, the identity matrix, and $\boldsymbol{\varepsilon}_t$ are serially uncorrelated iid residuals (white noise) with covariance matrix $\Sigma \equiv \mathrm{cov}(\boldsymbol{\varepsilon}_t)$. We allow the model order p to be infinite. Introducing the lag operator $\mathcal{L}$ so that $\mathcal{L}\boldsymbol{U}_t = \boldsymbol{U}_{t-1}$, $\mathcal{L}^2\boldsymbol{U}_t = \boldsymbol{U}_{t-2}$, etc., we can write (1) in the form

$$A(\mathcal{L}) \cdot \boldsymbol{U}_t = \boldsymbol{\varepsilon}_t \tag{2}$$

where the p th order square matrix polynomial $A(z)$ is defined to be $A(z) \equiv \sum_{k=0}^{p} A_k z^k$, with $A(0) = I$.

Covariance-stationarity requires that $A(z)$ exists and is *invertible* for all z on the unit disk $|z| \leq 1$ in the complex z-plane (Hamilton, 1994); a VAR model of the form (2) is described as *stable* if it satisfies this condition. For the finite order case, this requires that all roots of the characteristic polynomial $\det(A(z^{-1}))$ lie strictly inside the unit circle. The maximum modulus of the roots of the characteristic polynomial is the *spectral radius* of the VAR model, written $\rho(A)$. Intuitively, $\rho(A)$ determines how rapidly autocorrelation of the VAR decays with increasing lag time, and stability requires that $\rho(A) < 1$.

Since the VAR (2) is assumed covariance-stationary, by the Wold decomposition theorem (Hamilton, 1994) it may be written equivalently in VMA form as

$$\boldsymbol{U}_t = H(\mathcal{L}) \cdot \boldsymbol{\varepsilon}_t \tag{3}$$

where the *transfer function* $H(z)$ for the model is the rational matrix function defined by $H(z) \equiv A(z)^{-1}$. In general, the VMA representation will be of *infinite* order.

2.1. Time domain

We consider firstly *unconditional* G-causality. Suppose that $\boldsymbol{U}_t$ is decomposed into two jointly distributed, multivariate processes $\boldsymbol{U}_t = \begin{pmatrix} \boldsymbol{X}_t \\ \boldsymbol{Y}_t \end{pmatrix}$ with $\dim(\boldsymbol{X}) = k$ and $\dim(\boldsymbol{Y}) = l$, $k + l = n$. We wish to ascertain the causal effect of the variable $\boldsymbol{Y}$ on the variable $\boldsymbol{X}$; i.e., the G-causality $\mathcal{F}_{\boldsymbol{Y} \to \boldsymbol{X}}$.

We may decompose the autoregression (2) as

$$\begin{pmatrix} A_{xx}(\mathcal{L}) & A_{xy}(\mathcal{L}) \\ A_{yx}(\mathcal{L}) & A_{yy}(\mathcal{L}) \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{X}_t \\ \boldsymbol{Y}_t \end{pmatrix} = \begin{pmatrix} \boldsymbol{\varepsilon}_{x,t} \\ \boldsymbol{\varepsilon}_{y,t} \end{pmatrix} \tag{4}$$

with VMA representation

$$\begin{pmatrix} \boldsymbol{X}_t \\ \boldsymbol{Y}_t \end{pmatrix} = \begin{pmatrix} H_{xx}(\mathcal{L}) & H_{xy}(\mathcal{L}) \\ H_{yx}(\mathcal{L}) & H_{yy}(\mathcal{L}) \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{\varepsilon}_{x,t} \\ \boldsymbol{\varepsilon}_{y,t} \end{pmatrix} \tag{5}$$

[3] In all that follows, bold type indicates a vector quantity and upper-case type denotes either a matrix or a random variable, depending on context. Vectors are considered to be *column* vectors. The symbol $^{\mathrm{T}}$ indicates matrix transpose; an asterisk denotes the conjugate transpose of a (complex) matrix, and $\det(\cdot)$ denotes the determinant of a (square) matrix.

and residuals covariance matrix

$$\Sigma = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix} \tag{6}$$

Since the sub-process $\boldsymbol{X}_t$ is covariance-stationary, by Wold's theorem it will itself have an (in general infinite order) VMA representation

$$\boldsymbol{X}_t = H'_{xx}(\mathcal{L}) \cdot \boldsymbol{\varepsilon}'_{x,t} \tag{7}$$

with transfer function $H'_{xx}(z)$, serially uncorrelated residuals $\boldsymbol{\varepsilon}'_{x,t}$ and residuals covariance matrix $\Sigma'_{xx} \equiv \mathrm{cov}(\boldsymbol{\varepsilon}'_{x,t})$. We then have a corresponding VAR form

$$A'_{xx}(\mathcal{L}) \cdot \boldsymbol{X}_t = \boldsymbol{\varepsilon}'_{x,t} \tag{8}$$

with $A'_{xx}(z) \equiv H'_{xx}(z)^{-1}$. We refer to (8) as the *restricted regression*, as opposed to the *full* or *unrestricted regression* (4). Importantly, even if the original VAR (2) is of finite order, the restricted VAR (8) will in general be of *infinite* order.

The (unconditional) MVGC from $\boldsymbol{Y}$ to $\boldsymbol{X}$ in the time domain (Geweke, 1982) is then defined to be[4]:

$$\mathcal{F}_{\boldsymbol{Y}\to\boldsymbol{X}} \equiv \ln\left(\frac{\det(\mathrm{cov}(\boldsymbol{\varepsilon}'_{x,t}))}{\det(\mathrm{cov}(\boldsymbol{\varepsilon}_{x,t}))}\right) = \ln\left(\frac{\det(\Sigma'_{xx})}{\det(\Sigma_{xx})}\right) \tag{9}$$

The rationale behind (9) is as follows: the magnitude of the residuals $\boldsymbol{\varepsilon}_{x,t}$ of the full regression (4) indicates how well both $\boldsymbol{X}$ and $\boldsymbol{Y}$ together predict the future of $\boldsymbol{X}$, while the magnitude of the residuals $\boldsymbol{\varepsilon}'_{x,t}$ of the restricted regression (8) indicates how well $\boldsymbol{X}$ predicts its *own* future. Thus $\mathcal{F}_{\boldsymbol{Y}\to\boldsymbol{X}}$ may be considered as a measure of the extent to which $\boldsymbol{Y}$ helps predict the future of $\boldsymbol{X}$ over and above the degree to which $\boldsymbol{X}$ already predicts its own future. It is strictly in this sense that MVGC should be considered a "causal" measure. In sample, (9) has a simple interpretation as a likelihood ratio test statistic under the null hypothesis of zero causality, $H_0 : A_{xy,k} = 0$ for $k = 1, \ldots, p$ (Geweke, 1982).

We have previously shown that, for Gaussian processes, $\mathcal{F}_{\boldsymbol{Y}\to\boldsymbol{X}}$ is equivalent to the *transfer entropy* from $\boldsymbol{Y}$ to $\boldsymbol{X}$ (Barnett et al., 2009), a measure of the time-directed information flow from the process $\boldsymbol{Y}$ to the process $\boldsymbol{X}$ (Schreiber, 2000; Kaiser and Schreiber, 2002). This information-theoretic interpretation is significant, as it implies that G-causality may be considered an absolute quantity (measured in bits) so that comparison of causalities between different sets of variables is valid (although validity may be undermined by differences in statistical bias; see Section 4.1.2).

2.2. Frequency domain

We write $A(\lambda)$ for the Fourier transform of the regression coefficients A_k, which is just $A(z)$ evaluated at $z = e^{-i\lambda}$ (it should be clear from context and notation when we refer to the frequency-domain version). The frequency domain transfer function is then $H(\lambda) = A(\lambda)^{-1}$ and the (cross) spectral power density of the multivariate process $\boldsymbol{U}_t$ is given by

$$S(\lambda) = H(\lambda)\Sigma H(\lambda)^* \tag{10}$$

A principled formulation of G-causality in the frequency domain was developed originally by Geweke (1982) as follows: the power spectrum may be decomposed as:

$$S(\lambda) = \begin{pmatrix} S_{xx}(\lambda) & S_{xy}(\lambda) \\ S_{yx}(\lambda) & S_{yy}(\lambda) \end{pmatrix} \tag{11}$$

Then $S_{xx}(\lambda)$ is just the spectral density of $\boldsymbol{X}_t$—which by (7) is also equal to $H'(\lambda)\Sigma' H'(\lambda)^*$—and from (10) we have

$$\begin{aligned} S_{xx}(\lambda) = {} & H_{xx}(\lambda)\Sigma_{xx}H_{xx}(\lambda)^* + 2\,\Re e\{H_{xx}(\lambda)\Sigma_{xy}H_{xy}(\lambda)^*\} \\ & + H_{xy}(\lambda)\Sigma_{yy}H_{xy}(\lambda)^* \end{aligned} \tag{12}$$

Geweke then notes that in the case that $\Sigma_{xy} \equiv 0$, which may always be effected by a linear transformation of variables leaving $\mathcal{F}_{\boldsymbol{Y}\to\boldsymbol{X}}$ invariant (Barrett et al., 2010), (12) takes the simpler form

$$S_{xx}(\lambda) = H_{xx}(\lambda)\Sigma_{xx}H_{xx}(\lambda)^* + H_{xy}(\lambda)\Sigma_{yy}H_{xy}(\lambda)^* \tag{13}$$

whereby the power spectrum of $\boldsymbol{X}$ splits into an "intrinsic" term and a "causal" term. The (unconditional) spectral MVGC from $\boldsymbol{Y}$ to $\boldsymbol{X}$ is then defined as

$$\mathrm{f}_{\boldsymbol{Y}\to\boldsymbol{X}}(\lambda) \equiv \ln\left(\frac{\det(S_{xx}(\lambda))}{\det(H_{xx}(\lambda)\Sigma_{xx}H_{xx}(\lambda)^*)}\right) \tag{14}$$

or, in terms of untransformed variables (i.e., where $\Sigma_{xy} \neq 0$),

$$\mathrm{f}_{\boldsymbol{Y}\to\boldsymbol{X}}(\lambda) \equiv \ln\left(\frac{\det(S_{xx}(\lambda))}{\det(S_{xx}(\lambda) - H_{xy}(\lambda)\Sigma_{y|x}H_{xy}(\lambda)^*)}\right) \tag{15}$$

with $S_{xx}(\lambda)$ as in (12) and the partial residuals covariance matrix $\Sigma_{y|x}$ is defined to be $\Sigma_{yy} - \Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy}$. Geweke then establishes the fundamental decomposition of MVGC by frequency:

$$\frac{1}{\pi}\int_0^\pi \mathrm{f}_{\boldsymbol{Y}\to\boldsymbol{X}}(\lambda)\,\mathrm{d}\lambda \leq \mathcal{F}_{\boldsymbol{Y}\to\boldsymbol{X}} \tag{16}$$

with equality when $\det(A_{yy}(z) - \Sigma_{yx}\Sigma_{xx}^{-1}A_{xy}(z)) \neq 0$ on the unit disk $|z| \leq 1$. We note, following Geweke (1982), that this condition for equality is usually satisfied (in particular it is satisfied for the model we examine in Section 4).

2.3. The conditional case

Suppose that $\boldsymbol{U}_t$ decomposes into *three* jointly distributed multivariate processes $\boldsymbol{U}_t = \begin{pmatrix} \boldsymbol{X}_t \\ \boldsymbol{Y}_t \\ \boldsymbol{Z}_t \end{pmatrix}$ with $\dim(\boldsymbol{X}) = k$, $\dim(\boldsymbol{Y}) = l$ and $\dim(\boldsymbol{Z}) = m$, $k + l + m = n$. We now wish to calculate the causal effect of the variable $\boldsymbol{Y}$ on the variable $\boldsymbol{X}$, controlling for any common effects of $\boldsymbol{Z}$ on both $\boldsymbol{X}$ and $\boldsymbol{Y}$; i.e., the G-causality $\mathcal{F}_{\boldsymbol{Y}\to\boldsymbol{X}|\boldsymbol{Z}}$. Following Geweke (1984), consider the full regression

$$\begin{pmatrix} A_{xx}(\mathcal{L}) & A_{xy}(\mathcal{L}) & A_{xz}(\mathcal{L}) \\ A_{yx}(\mathcal{L}) & A_{yy}(\mathcal{L}) & A_{yz}(\mathcal{L}) \\ A_{zx}(\mathcal{L}) & A_{zy}(\mathcal{L}) & A_{zy}(\mathcal{L}) \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{X}_t \\ \boldsymbol{Y}_t \\ \boldsymbol{Z}_t \end{pmatrix} = \begin{pmatrix} \boldsymbol{\varepsilon}_{x,t} \\ \boldsymbol{\varepsilon}_{y,t} \\ \boldsymbol{\varepsilon}_{z,t} \end{pmatrix} \tag{17}$$

and the restricted regression

$$\begin{pmatrix} A'_{xx}(\mathcal{L}) & A'_{xz}(\mathcal{L}) \\ A'_{zx}(\mathcal{L}) & A'_{zz}(\mathcal{L}) \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{X}_t \\ \boldsymbol{Z}_t \end{pmatrix} = \begin{pmatrix} \boldsymbol{\varepsilon}'_{x,t} \\ \boldsymbol{\varepsilon}'_{z,t} \end{pmatrix} \tag{18}$$

The conditional MVGC (Geweke, 1984) from $\boldsymbol{Y}$ to $\boldsymbol{X}$ given $\boldsymbol{Z}$ in the time domain is then defined as:

$$\begin{aligned} \mathcal{F}_{\boldsymbol{Y}\to\boldsymbol{X}|\boldsymbol{Z}} &\equiv \ln\left(\frac{\det(\mathrm{cov}(\boldsymbol{\varepsilon}'_{x,t}))}{\det(\mathrm{cov}(\boldsymbol{\varepsilon}_{x,t}))}\right) \\ &= \ln\left(\frac{\det(\Sigma'_{xx})}{\det(\Sigma_{xx})}\right) \end{aligned} \tag{19}$$

with similar rationale as for the unconditional case: $\mathcal{F}_{\boldsymbol{Y}\to\boldsymbol{X}|\boldsymbol{Z}}$ is to be considered as a measure of the extent to which $\boldsymbol{Y}$ helps predict the future of $\boldsymbol{X}$ over and above the degree to which $\boldsymbol{X}$ and $\boldsymbol{Z}$ *together* already predict the future of $\boldsymbol{X}$. (Note that the transfer entropy equivalence (Barnett et al., 2009) carries through to this conditional case.)

The spectral conditional case is less straightforward; Geweke notes that, defining the new variables $\boldsymbol{X}_t^\dagger \equiv \boldsymbol{\varepsilon}'_{x,t}$, $\boldsymbol{Z}_t^\dagger \equiv \boldsymbol{\varepsilon}'_{z,t}$ (i.e.,

[4] $\det(\mathrm{cov}(\boldsymbol{\varepsilon}))$ is known as the "generalised variance" (Barrett et al., 2010) of the residuals $\boldsymbol{\varepsilon}$. For a full discussion as to why the generalised variance is to be preferred to the "total variance" $\mathrm{trace}(\mathrm{cov}(\boldsymbol{\varepsilon}))$ in the context of G-causality [*cf.* Ladroue et al. (2009)], please see Barrett et al. (2010).

the residuals of the restricted regression (18)) and setting $\boldsymbol{YZ}^\dagger \equiv \begin{pmatrix} \boldsymbol{Y}_t \\ \boldsymbol{Z}_t^\dagger \end{pmatrix}$, we have the identity

$$\mathcal{F}_{\boldsymbol{Y}\to\boldsymbol{X}|\boldsymbol{Z}} \equiv \mathcal{F}_{\boldsymbol{YZ}^\dagger\to\boldsymbol{X}^\dagger} \tag{20}$$

Thus in the time domain the conditional MVGC may be expressed as an *unconditional* MVGC in terms of new variables defined as the residuals of the restricted regression. The frequency domain conditional MVGC is accordingly defined as

$$\mathrm{f}_{\boldsymbol{Y}\to\boldsymbol{X}|\boldsymbol{Z}}(\lambda) \equiv \mathrm{f}_{\boldsymbol{YZ}^\dagger\to\boldsymbol{X}^\dagger}(\lambda) \tag{21}$$

and the spectral decomposition

$$\frac{1}{\pi}\int_0^\pi \mathrm{f}_{\boldsymbol{Y}\to\boldsymbol{X}|\boldsymbol{Z}}(\lambda)\,\mathrm{d}\lambda \le \mathcal{F}_{\boldsymbol{Y}\to\boldsymbol{X}|\boldsymbol{Z}} \tag{22}$$

again holds, with equality under a corresponding condition to that for (16).

2.4. Application to empirical data

To apply the above formalism to empirical data, suppose we are given covariance stationary multivariate time series data, and that we may assume the underlying generative process to be reasonably modelled as a stable VAR(p) (of unknown order) of the form (1). There are then several strategies available for obtaining estimates of the necessary quantities for calculation of (time or frequency domain, unconditional or conditional) MVGCs. Most straightforwardly, an appropriate empirical model order p is first determined by, for example, the Akaike or Bayesian information criterion, cross-validation, or other standard technique[5] (McQuarrie and Tsai, 1998). Regression coefficients may then be estimated by one of several standard procedures, such as OLS, solution of the Yule-Walker relations or Whittle's multivariate extension of Levinson–Durbin recursion (Hamilton, 1994; Whittle, 1963). Once known, residuals may be calculated directly from the data and estimates obtained for their covariance matrices. For frequency-domain MVGC, the transfer function may be calculated by Fourier transforming regression coefficients, from which spectral power densities are then easily calculated according to (10). This is the approach taken in the empirical study (Section 4) in this paper; spectral MVGCs are calculated as just described, and time-domain MVGCs by integration of the corresponding spectral causalities.[6] Alternative approaches, not considered here, include the "nonparametric" method proposed by Dhamala et al. (2008a,b), in which the transfer function and residuals covariance matrices are obtained by canonical factorisation of the spectral density matrix (Wilson, 1972), which may be estimated directly from time series data by Fourier or wavelet transform.

2.4.1. Significance testing

For the *univariate predictee* case [i.e., where $k \equiv \dim(\boldsymbol{X}) = 1$], since the residuals covariance matrices are simple variances, a standard F-test on the restricted and full regressions under the null hypothesis $H_0 : \mathcal{F}_{\boldsymbol{Y}\to\boldsymbol{X}|\boldsymbol{Z}} = 0$ (and assumptions of normality for the residuals) yields, in both the unconditional and conditional case

$$\frac{N-p(l+m+1)}{pl}[\exp(\hat{\mathcal{F}}_{\boldsymbol{Y}\to\boldsymbol{X}|\boldsymbol{Z}})-1] \sim F(pl, N-p(l+m+1)) \tag{23}$$

where $\hat{\mathcal{F}}_{\boldsymbol{Y}\to\boldsymbol{X}|\boldsymbol{Z}}$ is the maximum likelihood estimator for $\mathcal{F}_{\boldsymbol{Y}\to\boldsymbol{X}|\boldsymbol{Z}}$, N is the sample size (time series length) and $F(d_1, d_2)$ denotes the F-distribution with d_1, d_2 degrees of freedom. For the unconditional case, (23) holds with $m = 0$. Unfortunately, for the multivariate predictee case $k > 1$, there does not appear to be an equivalent result. Geweke (1982) states that, at least for the unconditional case,[7] $N\hat{\mathcal{F}}_{\boldsymbol{Y}\to\boldsymbol{X}}$ approaches a $\chi^2(pkl)$ distribution asymptotically for large sample size N, although it is not clear what constitutes a "sufficiently large" sample for the approximation to be useful (*cf.* Section 4.1.2 below). For spectral MVGC (both uni- and multivariate, unconditional and conditional) nothing appears to be known about the distribution of the maximum likelihood estimator. For these reasons, non-parametric methods such as permutation testing (Anderson and Robinson, 2001) are generally preferable in order to obtain an empirical null distribution for significance testing. In the models analysed in Section 4, we compare all three methods (F, χ^2, and permutation testing).

3. Invariance of MVGC under multivariate filtering

Fig. 1 provides an example of the result we will derive analytically in this section. The left panel shows frequency-domain G-causality from one variable to another, before and after lowpass filtering (right panel). It is apparent that the G-causality in the stopband (shaded area), although noisy due to the sample estimation procedure (*cf.* Section 3.1), is essentially unchanged by the filtering process. We now explain the theoretical basis for this possibly counterintuitive result.

Suppose given a multivariate discrete digital filter with rational transfer function $G(z) = P(z)^{-1}Q(z)$, where $Q(z) = \sum_{k=0}^{r} Q_k z^k$ and $P(z) = \sum_{l=0}^{s} P_l z^l$ are $n \times n$ square matrix polynomials,[8] normalised so that $P(0) = I$ (the identity matrix). The filter is of FIR type iff $P(z) \equiv I$; otherwise it is of IIR type. We demand, furthermore, that the filter be *stable* and *invertible*. Stability requires that $\det(P(z)) \neq 0$ on the unit disk $|z| \le 1$; i.e., that all poles of $G(z)$ lie outside the unit circle (Antoniou, 1993), while invertibility requires that the $Q(0)$ be invertible. Intuitively, a filter is stable if an impulse does not "blow up". Invertibility guarantees that a pure lag inverse filter exists.[9] We note that a FIR filter is always stable.

We indicate filter-transformed quantities by a tilde, so that for a multivariate time series $\boldsymbol{u}_t$ the filter action in the time domain may be represented as $\tilde{\boldsymbol{u}}_t = G(\mathcal{L}) \cdot \boldsymbol{u}_t$, or

$$\sum_{l=0}^{s} P_l \cdot \tilde{\boldsymbol{u}}_{t-l} = \sum_{k=0}^{r} Q_k \cdot \boldsymbol{u}_{t-k} \tag{24}$$

In practice it is frequently the case that the same filter is applied individually to each component of $\boldsymbol{u}_t$ with no cross terms; i.e., $G(z) = \mathrm{diag}(g(z))$ for a univariate discrete digital filter with transfer function $g(z)$ of the more familiar form

$$g(z) = \frac{q_0 + q_1 z + \cdots + q_r z^r}{p_0 + p_1 z + \cdots + p_s z^s} \tag{25}$$

[5] Strictly speaking, following our remarks in Section 2.1, model orders should be estimated *separately* for the full and restricted regressions; however this seems rarely implemented in practice.

[6] Note that this approach—at least in the *un* conditional case—avoids having to fit a separate restricted regression.

[7] It appears that Geweke's argument applies equally to the conditional time domain statistic, again implying a $\chi^2(pkl)$ asymptotic distribution for $N\hat{\mathcal{F}}_{\boldsymbol{Y}\to\boldsymbol{X}|\boldsymbol{Z}}$, although in his subsequent paper (Geweke, 1984) introducing multivariate conditional G-causality this is not mentioned.

[8] We take $G(z)$ as a rational matrix function of z rather than, as is more common in the signal processing literature, of z^{-1}. This is consistent with the usage of Section 2; again, in the time domain z may be replaced by the lag operator $\mathcal{L}$, and in the frequency domain by $e^{-i\lambda}$.

[9] Note that this is rather a stringent condition: for example the simple delay FIR filter $\tilde{\boldsymbol{u}}_t = \boldsymbol{u}_{t-1}$, although stable, is not invertible.

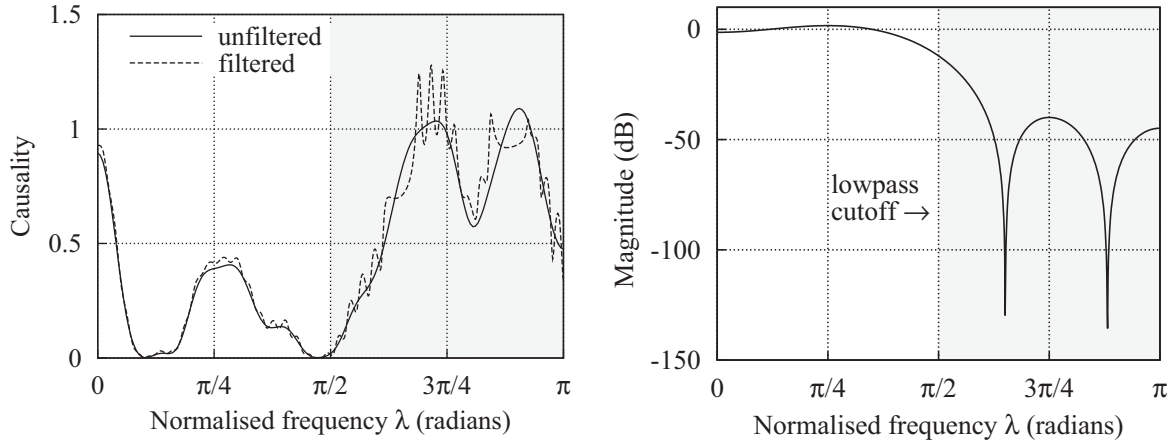

Fig. 1. Left-hand panel: frequency-domain (unconditional) G-causality from one variable to another, estimated in sample, before and after lowpass filtering. Time series data were generated from a simulated two variable VAR(30) process. Right-hand panel: frequency response of the 8th order least-squares linear-phase FIR filter. Shaded areas indicate the filter stop-band.

If now $\boldsymbol{U}_t$ is a covariance-stationary VAR(p) process as specified by (2) then the filtered process $\tilde{\boldsymbol{U}}_t \equiv G(\mathcal{L}) \cdot \boldsymbol{U}_t$ satisfies $A(z)G(z)^{-1} \cdot \tilde{\boldsymbol{U}}_t = \boldsymbol{\varepsilon}_t$. Thus it has the VAR representation

$$\tilde{\mathrm{A}}(\mathcal{L}) \cdot \tilde{\boldsymbol{U}}_t = \tilde{\boldsymbol{\varepsilon}}_t \tag{26}$$

with coefficients polynomial

$$\tilde{\mathrm{A}}(z) \equiv G(0)A(z)G(z)^{-1} \tag{27}$$

and residuals $\tilde{\boldsymbol{\varepsilon}}_t \equiv G(0) \cdot \boldsymbol{\varepsilon}_t$. Since the filter is assumed invertible, $G(0)^{-1}$ exists and thus $\tilde{A}(0) = I$ as required, while from the filter stability assumption it is clear that the VAR model (26) will be stable, since the zeros of det $(\tilde{A}(z))$ are the zeros of det$(A(z))$ together with the zeros of det$(P(z))$. Importantly, even if the unfiltered VAR (2) has finite model order, the filtered VAR (26) will in general have *infinite* model order; this may be seen from the presence of the $G(z)^{-1}$ term in the expression (27) for the filtered VAR coefficients polynomial.[10]

Starting with the unconditional case, given a decomposition $\boldsymbol{U}_t = \begin{pmatrix} \boldsymbol{X}_t \\ \boldsymbol{Y}_t \end{pmatrix}$ as before, we ask how the corresponding G-causality $\mathcal{F}_{\tilde{\boldsymbol{Y}} \to \tilde{\boldsymbol{X}}}$ transforms under the filter, where the filtered decomposition is

$$\begin{pmatrix} \tilde{\boldsymbol{X}}_t \\ \tilde{\boldsymbol{Y}}_t \end{pmatrix} = \begin{pmatrix} G_{xx}(\mathcal{L}) & G_{xy}(\mathcal{L}) \\ G_{yx}(\mathcal{L}) & G_{yy}(\mathcal{L}) \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{X}_t \\ \boldsymbol{Y}_t \end{pmatrix} \tag{28}$$

From (28) we see that as long as $G_{xy}(z) \equiv 0$—i.e., the filtered components of $\boldsymbol{X}$ do not depend on $\boldsymbol{Y}$—we have

$$\tilde{\boldsymbol{X}}_t = G_{xx}(\mathcal{L}) \cdot \boldsymbol{X}_t \tag{29}$$

whence from (8) the restricted filtered process has the VAR representation

$$\tilde{\mathrm{A}}'_{xx}(\mathcal{L}) \cdot \tilde{\boldsymbol{X}}_t = \tilde{\boldsymbol{\varepsilon}}'_{x,t} \tag{30}$$

with coefficients polynomial $\tilde{\mathrm{A}}'_{xx}(z) \equiv G_{xx}(0)A'_{xx}(z)G_{xx}(z)^{-1}$ and residuals $\tilde{\boldsymbol{\varepsilon}}'_{x,t} \equiv G_{xx}(0) \cdot \boldsymbol{\varepsilon}'_{x,t}$. Now the filtered full regression (26) has residuals covariance matrix $\tilde{\Sigma} = G(0)\Sigma G(0)^{\mathsf{T}}$, so that in particular $\tilde{\Sigma}_{xx} = G_{xx}(0)\Sigma_{xx}G_{xx}(0)^{\mathsf{T}}$, while the filtered restricted regression (30) has residuals covariance matrix $\tilde{\Sigma}'_{xx} = G_{xx}(0)\Sigma'_{xx}G_{xx}(0)^{\mathsf{T}}$. Now $G_{xy}(z) \equiv 0$ and $G(0)$ non-singular imply that $G_{xx}(0)$ is also non-singular. It follows immediately from (9)—the determinants det$(G_{xx}(0))$ factor out and cancel—that

$$\mathcal{F}_{\tilde{\boldsymbol{Y}} \to \tilde{\boldsymbol{X}}} = \mathcal{F}_{\boldsymbol{Y} \to \boldsymbol{X}} \tag{31}$$

so that G-causality in the time domain remains invariant[11] under any stable, invertible filter $G(z)$ with $G_{xy}(z) \equiv 0$.

In the frequency domain, $\tilde{H}(\lambda) = G(\lambda)H(\lambda)G(0)^{-1}$, so that $\tilde{S}(\lambda) = G(\lambda)S(\lambda)G(\lambda)^*$. Thus in particular, $\tilde{H}_{xy}(\lambda) = G_{xx}(\lambda)H_{xy}(\lambda)G_{yy}(0)^{-1}$ and $\tilde{S}_{xx}(\lambda) = G_{xx}(\lambda)S_{xx}(\lambda)G_{xx}(\lambda)^*$, while the partial residuals covariance matrix transforms as $\tilde{\Sigma}_{y|x} = G_{yy}(0)\Sigma_{y|x}G_{yy}(0)^{\mathsf{T}}$. Thus in (15) the determinants det$(G_{xx}(\lambda))$ factor out and cancel and we have at (almost[12]) all frequencies λ

$$\mathrm{f}_{\tilde{\boldsymbol{Y}} \to \tilde{\boldsymbol{X}}}(\lambda) = \mathrm{f}_{\boldsymbol{Y} \to \boldsymbol{X}}(\lambda) \tag{32}$$

Thus spectral G-causality demonstrates the same invariance as in the time domain.

It may be verified along similar lines that the invariance extends (in both time and frequency domain) to the conditional case, where we now require $G_{xy}(z) = G_{xz}(z) = G_{zy}(z) \equiv 0$. This result may be considered a generalisation of the invariance of MVGC under the group of (unlagged) linear transformations of variables given in Barrett et al. (2010, Section 4.2).

3.1. Non-invariance in practice

Given the theoretical invariance demonstrated above, how can we account for the disruptions of G-causality estimates by filtering that have been noted in simulations (e.g., Seth, 2010; Florin et al., 2010)? First, we emphasise that the invariance holds strictly for *s*tationary processes which may be reasonably modelled as VARs. If the data is to begin with non-stationary, then G-causality analysis is inappropriate[13] and the resulting (in any case spurious) results may well be altered by filtering. In fact it can be useful to apply filtering

[10] The exception is an all-pole IIR filter, in which case the order of the filtered VAR is the sum of the VAR order and the filter order; such filters are unusual in practice and are not considered here. In general, digital filtering preserves finite order in VARMA processes, but not in VAR processes.

[11] Note that for the trace version of MVGC (Ladroue et al., 2009) the invariance will *not* hold in general, since the trace of $G_{xx}(0)$ does not factor out.

[12] It is possible that $G_{xx}(\lambda)$ vanishes at some frequencies, at which the spectral G-causality becomes undefined. At worst, however, this may only occur at a finite, discrete set of frequencies. We note that this situation cannot arise if we impose the further restriction that the inverse filter also be stable; in fact it appears that this restriction is required to guarantee preservation of equality in (16).

[13] It is possible to define (and estimate) G-causalities for non-stationary time series if multiple (synchronised) realisations of a process are available (Ding et al., 2000). We do not address this case here.

to attain stationarity, for example by the use of notch filtering to remove line noise (*cf.* Section 4.1.3), or highpass filtering to remove low-frequency transients. Even the common preprocessing step of differencing may be viewed as a (stable, invertible) FIR filter.

Assuming a stationary VAR process, there are several interlinked reasons why filtering may corrupt G-causality estimates in sample:

(I) The first and most important is the increase in model order entailed by filtering. As noted, filtered processes will in general have *infinite* model order; yet in sample a finite model order must be chosen. This means that (i) any finite (estimated) model order approximation to the filtered VAR process will inevitably result in a poorer model fit than for the unfiltered process, and (ii) the increase in number of model parameters will result in poorer estimates of those parameters. For short time series (where high model order also decreases significantly the available time series length) these effects may make it virtually impossible to estimate the "real" (high) filtered model order parameters without overfitting. Standard model order selection procedures, along with noisy parameter estimation (Section 2.4) are thus likely to result in substantially sub-optimal estimated models and, consequently, poor causal estimates with an increased likelihood of type I and type II errors in significance testing. To aid intuition on this issue, recall that in practice highpass, lowpass, bandpass and notch filters are often applied for which the frequency response[14] is very close to zero in the stop band, where filtered series will consequently have power spectra very close to zero. Further, for high-order filters (and particularly IIR filters) there may be steep roll-off on the edges of the stop band (e.g., elliptic filters) and/or broad, flat spectra in the pass band (e.g., Butterworth filters). Filtering thus "distorts" the power spectrum of the process so that filtered data will need to be modelled by a high order VAR to capture the detail in the modified spectrum.

(II) A related cause of error is that, even though in theory a stable filter acting on a stable VAR process yields another stable VAR process, filtering may increase the likelihood that empirical estimates of VAR parameters yield unstable or near-unstable models. The reason is again poor parameter estimation due to increased model order as discussed above. Furthermore, Eq. (27) shows that this effect is likely to be exacerbated in the case of IIR filters for which the poles of the filter transfer function lie closer to the unit circle than the poles of the VAR transfer function, effectively increasing the spectral radius for the filtered process. Note that an unstable estimated model precludes any further causal analysis.

(III) A final source of error is the potential appearance of numerical instabilities in the causal estimation procedure following filtering. In the frequency domain, since $\tilde{S}(\lambda) = G(\lambda)S(\lambda)G(\lambda)^*$ we have $\det(\tilde{S}(\lambda)) = \det(G(\lambda))^2 \cdot \det(\tilde{S}(\lambda))$ and from (14) or (15) we see that as the filter response approaches zero at some frequency λ, the spectral MVGC $\mathrm{f}_{\boldsymbol{Y}\to\boldsymbol{X}}(\lambda)$ becomes singular. In sample these relationships will not be exact, but may nevertheless result in numerical instabilities when calculating causalities from empirical data, in both the time and frequency domains. To see how such instabilities might occur in practice, consider the autoregressions necessary for calculation of both time and frequency domain G-causality. For a covariance-stationary VAR(p) as in (2) it may be shown by a Yule-Walker, OLS or equivalent procedure that the regression coefficients A_k can be expressed in terms of the *autocovariance sequence* $\Gamma_k \equiv \mathrm{cov}(\boldsymbol{U}_t, \boldsymbol{U}_{t-k})$, $k = \ldots, -2, -1, 0, 1, 2, \ldots$ of the process $\boldsymbol{U}_t$ as

$$(A_1 \quad A_2 \quad \ldots \quad A_p) = -(\Gamma_1 \quad \Gamma_2 \quad \ldots \quad \Gamma_p) \times \begin{pmatrix} \Gamma_0 & \Gamma_1 & \ldots & \Gamma_{p-1} \\ \Gamma_{-1} & \Gamma_0 & \ldots & \Gamma_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_{-(p-1)} & \Gamma_{-(p-2)} & \ldots & \Gamma_0 \end{pmatrix}^{-1} \tag{33}$$

For empirical data, computation of (33), with the Γ_k replaced by their sample estimates, is implicit in any standard[15] solution of the regression (1). The power spectrum of a covariance-stationary process $\boldsymbol{U}_t$ is (by definition) the Fourier transform of the autocovariance sequence: $S(\lambda) \equiv \sum_{k=-\infty}^{\infty} \Gamma_k e^{-i\lambda k}$. Suppose that $\boldsymbol{U}_t$ has been pre-filtered such that at some frequency λ, we have $S(\lambda) \approx 0$. Then we see that *near-colinearities* arise among the Γ_k with the consequence that the matrix inversion in (33) may become ill-conditioned, and estimation of both time and frequency domain G-causalities unreliable. One might consider alleviating such instabilities by adding (minimal) white noise; however, simulations (not reported) indicate that even very low level added noise in conjunction with severe filtering introduces artefacts which skew causal estimation.

Looking ahead, our simulation results (Section 4) show that degraded parameter estimation due to increased model order is the principal cause of poor G-causality estimation in sample. We expect this effect to be very general across all applications of filtering to stationary VAR processes. This effect also accounts for the observed increase in the incidence of unstable estimated models (Fig. 5). We note that for the IIR filter used, the poles lie further from the unit circle than the poles of the VAR transfer function (Fig. 4), so that there is no (theoretical) increase in spectral radius. Nor was there evidence of numerical instability in our simulations, due most likely to the low order VAR and comparatively low order of the filters used; it is not difficult, though, to construct examples where this effect is evident.

3.2. Band-limited MVGC

As mentioned, a common application of filtering is to restrict analysis to frequency ranges of prior interest. Our analysis indicates that this strategy is inappropriate for G-causality. We now suggest an alternative method for analysing G-causality in specific frequency bands, in both time and frequency domains. In the frequency domain the solution is trivial: we simply disregard values of $\mathrm{f}_{\boldsymbol{Y}\to\boldsymbol{X}}(\lambda)$ or $\mathrm{f}_{\boldsymbol{Y}\to\boldsymbol{X}|\boldsymbol{Z}}(\lambda)$ at frequencies λ outside the desired range. In the time domain, suppose that $\mathcal{B} \subset [0, \pi]$, the frequency band of prior interest, is a (not necessarily connected) measurable subset of the full (normalised) frequency range; that is, we wish to suppress causal contributions at frequencies lying outside $\mathcal{B}$. The spectral decomposition relations (16) and (22) may be viewed as averaging spectral causality over the full range frequency range $\lambda \in [0, \pi]$. Therefore causal contributions within the desired pass band $\mathcal{B}$ are given simply by the average (unfiltered) spectral causality over $\mathcal{B}$. Accordingly, we define

$$\mathcal{F}_{\boldsymbol{X}\to\boldsymbol{Y}}(\mathcal{B}) \equiv \frac{1}{\mu(\mathcal{B})} \int_{\mathcal{B}} \mathrm{f}_{\boldsymbol{Y}\to\boldsymbol{X}}(\lambda)\, \mathrm{d}\lambda \tag{34}$$

[14] For a multivariate filter with transfer function $G(z)$, we measure the magnitude of the frequency response (i.e., the gain) by $||\, G(\lambda)\, ||$, where $||\cdot||$ denotes some matrix norm.

[15] An exception is the nonparametric approach of Dhamala et al. (2008a,b) mentioned in Section 2.4. It is not clear how accuracy and numerical stability of this technique compare with more conventional regression-based methods, particularly for short time series.

$$\mathcal{F}_{\mathbf{X}\to\mathbf{Y}|\mathbf{Z}}(\mathcal{B}) \equiv \frac{1}{\mu(\mathcal{B})}\int_{\mathcal{B}} \mathrm{f}_{\mathbf{Y}\to\mathbf{X}|\mathbf{Z}}(\lambda)\,\mathrm{d}\lambda \tag{35}$$

where $\mu(\mathcal{B}) \equiv \int_{\mathcal{B}} \mathrm{d}\lambda$ is the measure (length) of $\mathcal{B}$. We term $\mathcal{F}_{\mathbf{X}\to\mathbf{Y}}(\mathcal{B})$, [resp. $\mathcal{F}_{\mathbf{X}\to\mathbf{Y}|\mathbf{Z}}(\mathcal{B})$] the unconditional [resp. conditional] *band-limited* multivariate G-causality over the frequency band(s) specified by $\mathcal{B}$. As with spectral G-causality, there is no known empirical null distribution for $\mathcal{F}_{\mathbf{X}\to\mathbf{Y}}(\mathcal{B})$ or $\mathcal{F}_{\mathbf{X}\to\mathbf{Y}|\mathbf{Z}}(\mathcal{B})$, so for significance testing non-parametric methods such as permutation testing should be employed (*cf.* Section 2.4.1).

4. A minimal example

In this section, we introduce a simple VAR model for which G-causalities can be analytically derived.[16] We use this model to explore the empirical issues described above. Consider the two-variable VAR(1) model

$$X_t = aX_{t-1} + cY_{t-1} + \varepsilon_{x,t} \tag{36}$$

$$Y_t = bY_{t-1} + \varepsilon_{y,t} \tag{37}$$

with uncorrelated, unit variance residuals. It is stable iff $|a| < 1$ and $|b| < 1$. This is the simplest system with non-trivial G-causalities; we therefore refer to it as the "minimal VAR(1)".

In the previous notation we have $\Sigma = I$, and $A_1 = -\begin{pmatrix} a & c \\ 0 & b \end{pmatrix}$, so that

$$A(z) = \begin{pmatrix} 1-az & -cz \\ 0 & 1-bz \end{pmatrix} \tag{38}$$

and

$$H(z) = \frac{1}{(1-az)(1-bz)}\begin{pmatrix} 1-bz & cz \\ 0 & 1-az \end{pmatrix} \tag{39}$$

By (10), since $\Sigma = I$, we have

$$S(z) = H(z)H(z)^* = \frac{1}{|1-az|^2|1-bz|^2} \times \begin{pmatrix} |1-bz|^2 + c^2|z|^2 & cz(1-a\bar{z}) \\ c\bar{z}(1-az) & |1-az|^2 \end{pmatrix} \tag{40}$$

and setting $z = e^{-i\lambda}$, from (14) we may derive

$$\mathrm{f}_{Y\to X}(\lambda) = \ln\left(1 + \frac{c^2}{1-2b\cos\lambda + b^2}\right) \tag{41}$$

We see that $\mathrm{f}_{Y\to X}(\lambda)$ does not depend on the autoregressive coefficient a of X_t and is monotonic decreasing (resp. increasing) according as $b > 0$ (resp. $b < 0$). From (16) (it may be verified that the condition for equality is trivially satisfied) the time domain MVGC is given by[17]

$$\mathcal{F}_{Y\to X} = \frac{1}{\pi}\int_0^\pi \ln\left(1 + \frac{c^2}{1-2b\cos\lambda + b^2}\right)\mathrm{d}\lambda \tag{42}$$

$$= \ln\left(\frac{1+b^2+c^2+\sqrt{(1+b^2+c^2)^2 - 4b^2}}{2}\right) \tag{43}$$

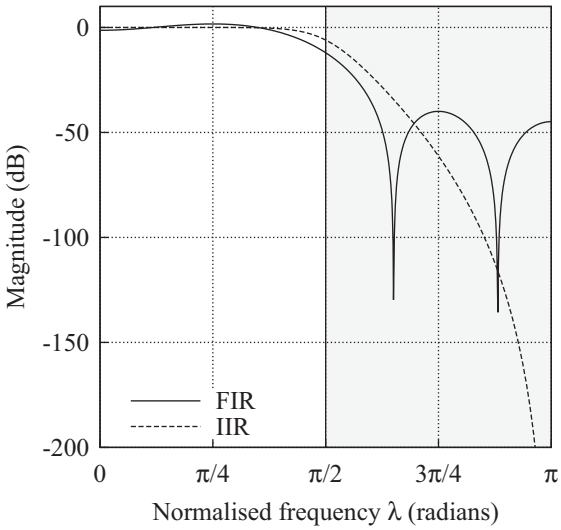

Fig. 2. Lowpass filter frequency response: the FIR (solid line) filter is an order 8 linear-phase least-squares, the IIR filter (dashed line) an order 4 Butterworth, both applied in forward and reverse directions for zero-phase filtering. Cutoff (normalised) frequency for both filters is at $\pi/2$ radians.

4.1. Experiments

Two types of (stable, invertible) digital filters were tested: a FIR linear-phase least-squares filter of order 8, and an IIR Butterworth filter of order 4 (Antoniou, 1993). Both filters were lowpass, with cutoff at normalised[18] frequency $\lambda = \pi/2$ radians. For consistency with standard practice, data was filtered in both the forward and reverse directions for zero-phase filtering[19]; note that this has the effect of squaring the transfer function, thus doubling the filter order. All filters were stable and invertible, with stable inverse filters (*cf.* our remarks in Section 3). Fig. 2 shows the frequency response in dB for both filter types. Note that neither filter is all-pole (*cf.* Section 3) so the filtered VAR will thus always be of infinite order.

We performed several experiments using the minimal VAR(1) specified by (36) and (37) to investigate the effects of filtering. In all experiments the system was simulated with parameters $a = 0.3$, $b = -0.8$ and $c = c_{sig} \approx 0.2104$ calculated from (43) so that $\mathcal{F}_{Y\to X} = 0.1$ for a "significant causality" model. For a "null model" (no causality) c was set to zero so that $\mathcal{F}_{Y\to X} = 0$. The theoretical spectral G-causality (41) for the significant causality model is plotted along with components of the power spectrum of the process in Fig. 3. We see that $\mathrm{f}_{Y\to X}(\lambda)$ peaks as the frequency approaches π radians, so that the lowpass filter (Fig. 2) suppresses power in a broad band where spectral causality is highest.

To illustrate clearly finite-sample effects, comparatively short stationary time series of length $N = 2^9 = 512$ were generated.[20] In all experiments regression coefficients were calculated by a standard OLS procedure, via QR decomposition of the stacked, lagged series. It is during this stage of computation that numerical instabilities as discussed in Section 3.1 under point III, might arise. As noted, this did not occur during any of our experiments. Frequency-domain causalities were then calculated according to (15) at a frequency resolution of $2^{10} = 1024$ over the normalised frequency range $\lambda \in [0,$

[16] As far as we are aware, this analysis is the first G-causality example in the literature to be solved in full analytically.

[17] The definite integral may be evaluated from its partial derivative with respect to c, which is integrable over $\lambda \in [0, \pi]$ by elementary methods; $\mathcal{F}_{Y\to X}$ may then be calculated simply as an indefinite integral over c, with the initial condition that $\mathcal{F}_{Y\to X} = 0$ at $c = 0$.

[18] Frequencies are normalised throughout, so that π radians corresponds to the Nyqvist frequency $f/2$ where f is the sampling frequency in Hz.

[19] Experiments (not included) suggest that zero-phase filtering in fact has minimal qualitative impact on results.

[20] Powers of two were used to maximise efficiency of the FFT (Fast Fourier Transform) algorithm used in spectral calculations.

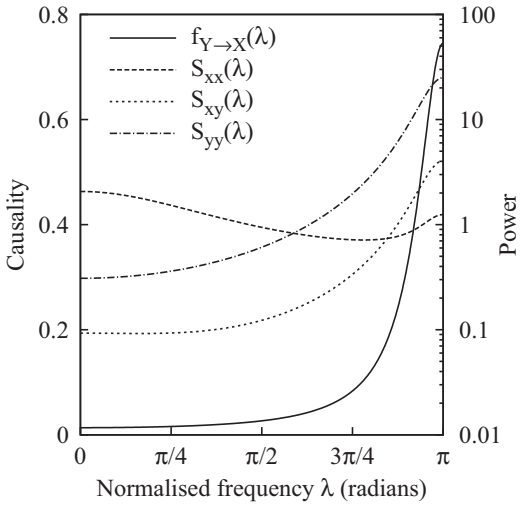

Fig. 3. Spectral G-causality $f_{Y\to X}(\lambda)$ of (41) (left axis) and power spectra $S(\lambda)$ calculated from (40) (right axis, logscale) of the minimal VAR(1) with $a = 0.3$, $b = -0.8$ and $c = c_{sig} \approx 0.2104$, plotted against normalised frequency λ.

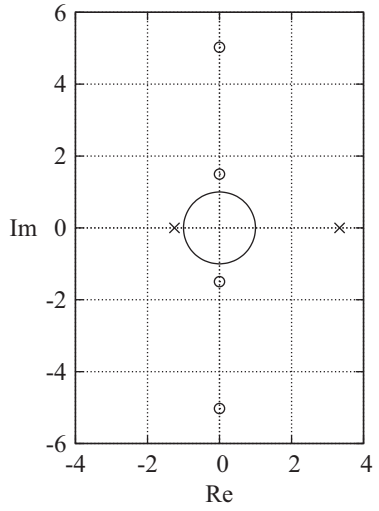

Fig. 4. Poles of the minimal VAR transfer function (crosses) and the 4th order Butterworth IIR filter (circles) drawn in the complex z-plane along with the unit circle.

π]. Finally, time-domain causalities were calculated by numerical quadrature of spectral causality according to (16) (*cf.* Section 2.4). Again we note that for the IIR filter, since all poles lie further from the unit circle than the poles $1/a$, $1/b$ of the transfer function (39) (Fig. 4), the potential confounding effect of increased spectral radius does not arise here (*cf.* Section 3.1).

4.1.1. Model order estimation and stability of estimated VAR for sample filtered process

We first examine model order estimation for the unfiltered and filtered time series. This involves calculating sample estimates $\hat{A}_1$, …, $\hat{A}_p$ of the regression coefficient matrices for a range of model orders p. Optimal model orders are then computed by various methods which balance model complexity against model fit (see below). Before this, it is first important to test whether estimated coefficients will in general define stable VAR s as model order increases. As noted, we can rule out unstable VAR estimates arising from the location of the filter poles. However, as model order increases, more parameters need to be estimated given the same data, which will eventually lead to unstable VAR estimates, for both unfiltered and filtered time series.

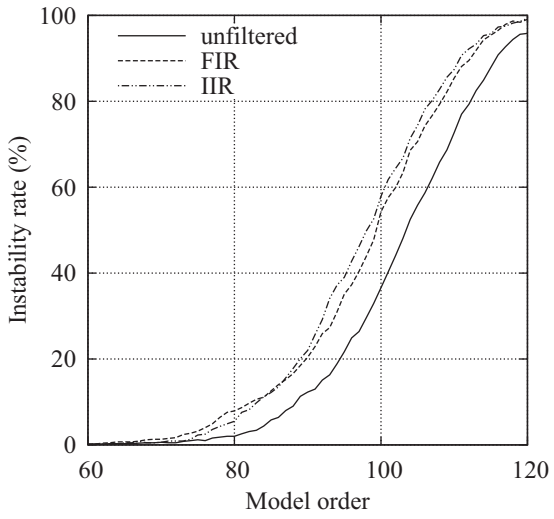

Fig. 5. Percentage of unstable VAR estimates plotted against model order for filtered and unfiltered series of length 2^9, for the minimal VAR(1) with $c = 0$.

To test this, we ran 1000 trial simulations of length $N = 2^9$ of the minimal VAR(1), for model orders ranging from 1 to 120. The trials were repeated with and without filtering, both for the significant $c = c_{sig}$ and null $c = 0$ systems. Regression coefficients for each trial were estimated as described above, and stability assessed via calculation of the spectral radius $\rho(\tilde{A})$ of the model (Section 2). Results for $c = 0$ (there was little difference for $c = c_{sig}$) are displayed in Fig. 5. The percentage of unstable VAR estimates increase sharply from order $p = 60$, reaching close to 100% by $p = 120$. As anticipated, instability rates are higher for filtered data given the increased difficulty of fitting VAR models following filtering (Section 3.1).

Three techniques were tested to assess an appropriate order for a VAR(p) model of the filtered (and unfiltered) data: the Akaike and Bayesian error criteria (AIC and BIC) and a cross-validation (CV) procedure (McQuarrie and Tsai, 1998). All are based on the maximised value of the log-likelihood function which, for an estimated VAR(p) model of the form (1), is given (up to a constant depending only on the sample size N, which remains fixed in our experiment) by

$$L \equiv -\frac{N}{2} \ln \det(\hat{\Sigma}) \tag{44}$$

where $\hat{\Sigma}$ is the sample estimate of the residuals covariance matrix. Then

$$\mathrm{AIC} = -2L + \frac{N(N+d)}{N-d-2} \tag{45}$$

$$\mathrm{BIC} = -2L + d \ln N \tag{46}$$

where $d \equiv pn^2$ is the number of parameters in the model. Optimal model orders are specified by the AIC/BIC reaching a minimum. We used a version (45) of the AIC which incorporates a second-order small sample correction as proposed by Hurvich and Tsai (1989), without which the AIC frequently failed to attain a minimum. For the cross-validation, VAR coefficients were estimated for training (in-sample) data and residuals calculated for independently simulated validation (out-of-sample) data of the same number of time steps. The CV model order estimation criterion was then to maximise the log-likelihood calculated from the out-of-sample residuals.[21] Note that for real-world data without the benefit of a

[21] Minimising the mean squared out-of-sample residuals gave virtually identical results.

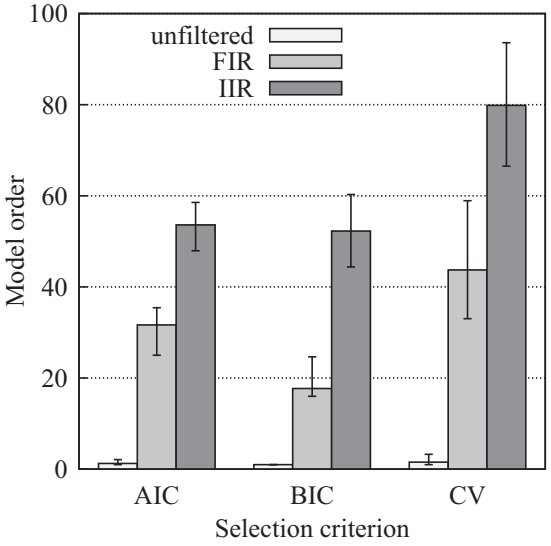

Fig. 6. Mean estimated optimal model order for unfiltered and filtered time series of length 2^9, for the minimal VAR(1) with $c = 0$ and various model order selection criteria. Error bars indicate 95% confidence intervals.

known generative process, CV estimates may be derived by standard techniques involving partitioning of the data into training and validation subsets.

For 1000 realisations of both the filtered and unfiltered minimal VAR(1) the various estimation criteria were calculated for model orders ranging from $p = 1$ to 120; unstable VAR estimates were discarded, as were cases where the criterion failed to achieve a minimum before the upper limit $p = 120$. Results for $c = 0$ (again, there was little difference for $c = c_{sig}$), showing the mean estimated optimal model order and 95% confidence intervals are displayed in Fig. 6. As expected, all estimation methods agreed on model order $p \approx 1$ for unfiltered data. For filtered data there is substantial disagreement between selection criteria; AIC and BIC tend to select model orders roughly half that of CV (interestingly, AIC and BIC are in agreement for the IIR but not the FIR filter). The cross-validation estimates are arguably the most reliable in principle (although in practice they require more data), yielding optimal mean values of $p \approx 44$ for the FIR filter and $p \approx 80$ for the IIR filter. Although the precise choice of model order does not affect our qualitative conclusions (see below), these values are clearly very much in excess of the actual unfiltered model order $p = 1$.

4.1.2. Statistical bias and significance testing

We next examined how filtering affected bias and significance testing of the G-causality test statistic. We performed another 1000 minimal VAR(1) simulations of length $N = 2^9$, filtered and unfiltered, for both the null and significant causality models, over a range of model orders. Again, unstable VAR estimates were discarded. Empirical distributions of $\mathcal{F}_{Y\to X}$ were calculated by Kaplan–Meier estimation (Kaplan and Meier, 1958), along with the F and asymptotic χ^2 null distributions. Results are displayed in Fig. 7.

We note firstly that, since the G-causality statistic $\mathcal{F}_{Y\to X}$ is positive, the corresponding finite-sample statistic $\mathcal{F}_{Y\to X}$ will suffer from systematic positive bias; thus, even for the null model, for which $\mathcal{F}_{Y\to X} = 0$, the sample statistic will generally be >0. The bias is seen clearly in Fig. 7; it increases strongly with model order, but is not substantially affected by filtering. One implication of this is that comparison of estimates from limited-sample data may be highly misleading, since the bias may vary between estimates (*cf.* Section 2.1). Note, however, that bias does not in itself affect hypothesis testing (Section 2.4.1 and below) since it will be reflected in the null distributions, both theoretical and empirically derived. For this reason, here we do not attempt to debias the G-causality test statistic (see e.g., Geweke, 1984, Section 4; Efron, 1982).

We now examine the effects of model order and filtering on statistical significance testing. First, for our small sample size it is not clear how accurately the asymptotic χ^2 or the F-distribution of Section 2.4.1 will approximate the distribution of $\mathcal{F}_{Y\to X}$ under the null hypothesis of zero causality (recall that the latter test is not available for a multivariate predictee variable). For the null model (Fig. 7, left column), sampled causalities *above* the critical lines indicate Type I errors (false positives) at 5% significance according to the corresponding null distribution (F, χ^2, or empirical,[22]) while for the significant causality model (right column), causalities *below* the critical lines indicate Type II errors (false negatives). The figure shows that the χ^2 null distribution substantially underestimates the critical value of $\mathcal{F}_{Y\to X}$ at 5% significance, while the F-distribution slightly overestimates it. These discrepancies increase strongly with model order, rather than with filtering *per se*; i.e., filtering impacts on significance testing via the increase in (estimated) model order, as indicated by the thick horizontal lines on the x-axes of Fig. 7: for the χ^2 distribution filtering effectively results in a sharp increase in Type I errors, while for the F-distribution, the picture is reversed, with a sharp increase in Type II errors (see also Fig. 8).

To gauge the effects of filtering on significance testing under a more realistic scenario where a null model is not known in advance, we also performed permutation tests. For both the significant and null VAR(1) models, for each of the 1000 trial sequences, 500 random permutation sequences were generated from the trial sequence by the method of Freedman and Lane (Anderson and Robinson, 2001); these were used to simulate an empirical null distribution which was then employed to test for significance at p-value 0.05. Fig. 8 shows Type I (left column) and Type II (right column) errors generated by permutation, χ^2 and F-distributions, for unfiltered data, and for filtered data using both FIR and IIR filters. The top row shows results using optimal model orders; middle and bottom rows use model orders of 16 and 4, respectively (see below). Considering first the top row (optimal model orders): For Type I errors, without filtering all null distributions give the expected error rate of ≈0.05. Consistent with Fig. 7, after filtering Type I error rates are generally higher both for the permutation-test empirical distribution and, especially, for the χ^2 distribution. The effect is stronger for the IIR filter, reflecting the higher model order. For Type II errors, without filtering errors are negligible under all null distributions. After filtering, Type II error rates are increased, for the permutation-test empirical distribution and, especially, for the F-distribution. Again, the effect is stronger for the IIR filter. In summary, results indicate an increase in both Type I and Type II errors under filtering, mostly attributable to an increase in estimated model order. In addition they support concerns (*cf.* Section 2.4.1) on the use of the theoretical null distributions for short time series.

In many applications, researchers may use model orders based on prior knowledge of the data generating process (the underlying mechanism) rather than on formal model order selection criteria such as AIC and BIC, especially if the latter are specifying excessively high model orders or not reaching a minimum. The middle and bottom rows show results from repeating the above experiments using fixed model orders of 16 (reflecting an "informed guess") and 4 (reflecting a drastic underestimation), respectively. For model

[22] Because the empirical null distribution is calculated from the *known* null model, the Type I error probability at p-value = α will always be exactly α by design. Thus in the left column the empirical null distribution critical lines coincide with the upper 95% confidence limit.

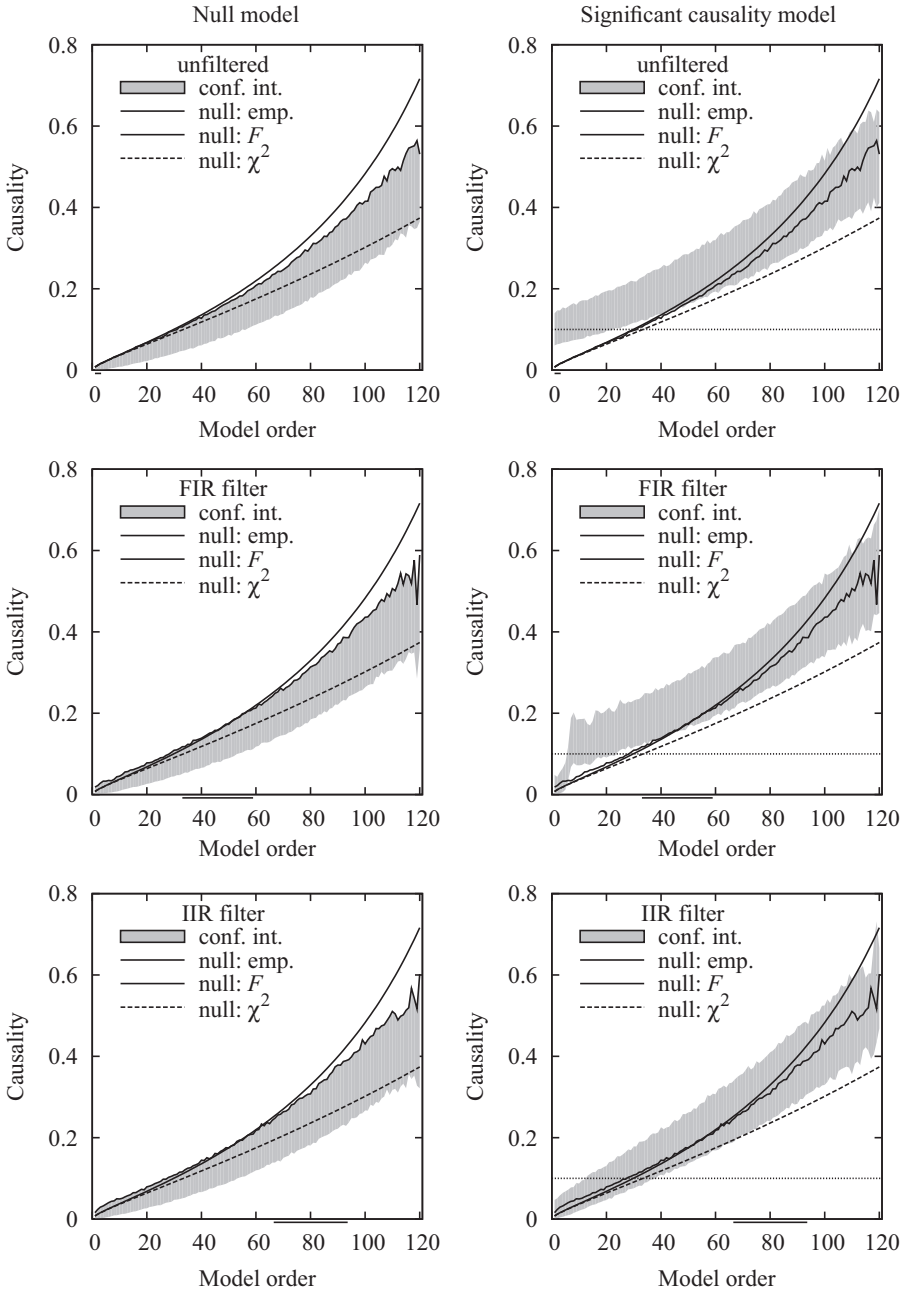

Fig. 7. Sample G-causality distributions $\mathcal{F}_{Y\to X}$ plotted against model order for filtered and unfiltered data, for the null model (left column) and significant causality model (right column). Shaded areas denote 95% confidence ranges for the $\mathcal{F}_{Y\to X}$ distribution. Bold solid lines indicate critical values for $\mathcal{F}_{Y\to X}$ at a 5% significance level (i.e., p-value = 0.05) under the null hypothesis $\mathcal{F}_{Y\to X} = 0$, calculated from the empirical null distribution; i.e., the sample distribution of $\mathcal{F}_{Y\to X}$ for the null model. Normal solid and dashed lines denote critical values at 5% significance for the theoretical F- and asymptotic χ^2 null distributions, respectively. The thick horizontal bars on the x-axis indicate 95% confidence intervals for the estimated optimum model order. For the significant causality model the horizontal line indicates the actual causality $\mathcal{F}_{Y\to X} = 0.1$.

order 16, for unfiltered data, Type I error rates are slightly higher as compared to the corresponding optimal model order results; Type II error rates remain negligible. For the filtered data, results show a complex pattern. Type I error rates are similar (as compared to optimal model order) for both the permutation test and F-distributions following FIR filtering, and are slightly higher under IIR filtering. For the χ^2 distribution, Type I errors are actually lower when using a model order 16 as compared to optimal, for both FIR and IIR filtering. Type II error rates are negligible under FIR filtering for all distributions (lower than the optimal model order), and are high and roughly equal under IIR filtering (corresponding to lower error for F, higher for χ^2, and similar for permutation distributions, as compared to the optimal model order). Summarising, an informed guess (underestimate) of model order performs well and in fact shows greater invariance with respect to the selected null distribution (F, χ^2, or permutation); these properties may reflect a tradeoff between "poor estimation of a good model" (at the estimated optimal model order), against "good estimation of a poor model" (at the arbitrarily chosen lower model order).

Results using model order of 4 (Fig. 8, bottom row), reflect a more drastic underestimation of model order; such low model orders are however often employed in practical applications in neuroscience (Bressler and Seth, 2011; Ding et al., 2006). For the filtered data, both Type I and Type II errors rates are now substantially higher, reflecting a poor model for the (known to be high order) filtered VAR. Interestingly, there is now very little

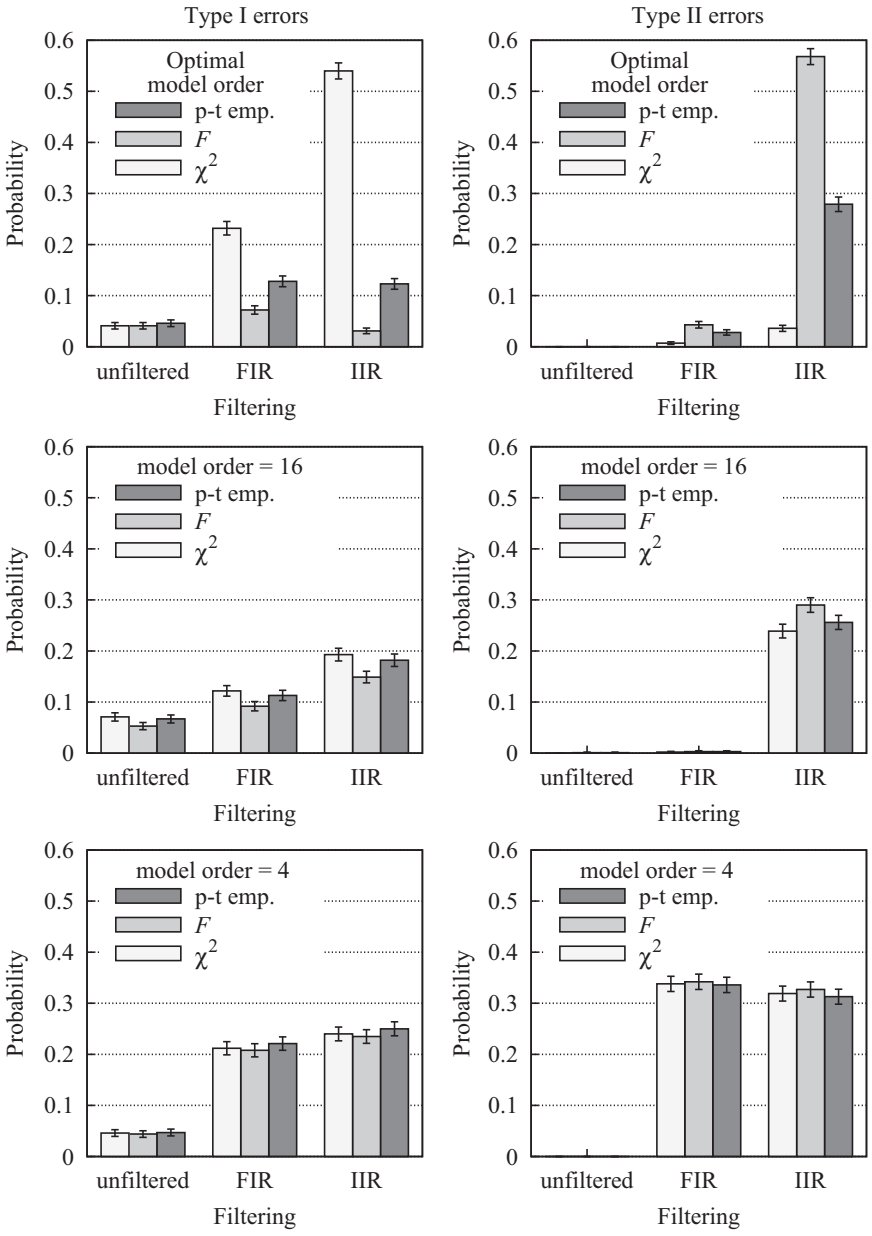

Fig. 8. Significance testing: probabilities of Type I errors (false positives, left) and Type II errors (false negatives, right) at 5% significance level for the different filter types, estimated from the permutation-test empirical ("p-t emp."), and F- and χ^2 null distributions. Error bars indicate standard errors. Top row: model orders are estimated optimal values: 1 for unfiltered, 44 for FIR and 80 for IIR filters (*cf.* Fig. 6). Middle row: model orders all set to 16. Bottom row: model orders all set to 4.

difference between results obtained using the various null distributions.

4.1.3. Eliminating line noise by notch filtering

As mentioned, G-causality is (theoretically) invariant to filtering only for stationary processes, leaving open the possibility that filtering may be useful for achieving stationarity, and/or for reducing model order of a nearly non-stationary process. To illustrate this usage, we examine notch filtering of time series data contaminated by fixed-frequency harmonic components (e.g., 50/60 Hz electrical line noise). We simulated line noise by adding a fixed sinusoidal signal of amplitude 1 and normalised frequency $\pi/4$ radians to both the X and Y component time series generated from the minimal VAR(1). Data was then filtered (in both forward and reverse directions) by two types of digital notch filters: a FIR linear-phase least-squares filter of order 64 (a high order was needed in order to alleviate ripple in the passband and curtail power sufficiently in the stopband), and an IIR Butterworth filter of order 4, both with a notch of width 0.02 radians (Fig. 9). Although strictly speaking the process with an added sinusoid is non-stationary, in finite sample it may be approximated and modelled as a VAR(p). As in Section 4.1.1 we estimated model order using AIC, BIC and CV in 1000 trials of filtered and unfiltered time series of length 2^9. Selected model orders (for the null model $c = 0$) are shown in Fig. 10. As expected, addition of line noise increases the model order considerably (although again with substantial disagreement among the selection criteria). Notch filters (both FIR and IIR) effectively reduce the model order to close to the value 1 of the original noiseless process.

Fig. 11 plots sampled causalities against model order (*cf.* Fig. 7), and Fig. 12 plots Type I and Type II error rates, for contaminated and filtered data (*cf.* Fig. 8). For unfiltered (i.e., contaminated) data, both the theoretical F- and, more markedly, the asymptotic χ^2 null distribution greatly underestimate the critical 5% significance level. In contrast to the filtered data in Fig. 7, this effect is stronger at

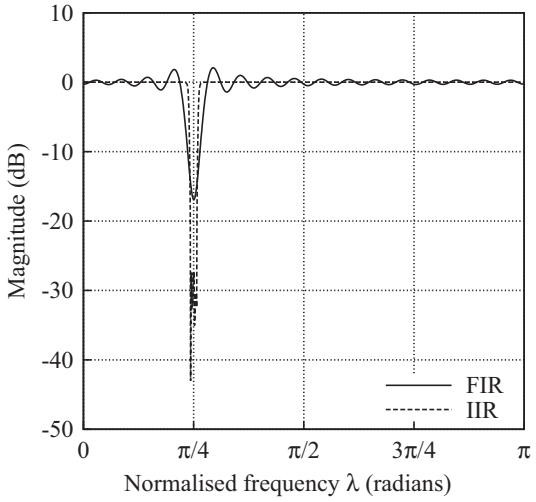

Fig. 9. Notch filter frequency response: the FIR (solid line) filter is an order 64 linear-phase least-squares, the IIR filter (dashed line) an order 4 Butterworth, applied in forward and reverse directions. Both have a notch of width 0.2 radians centred at $\pi/4$ radians.

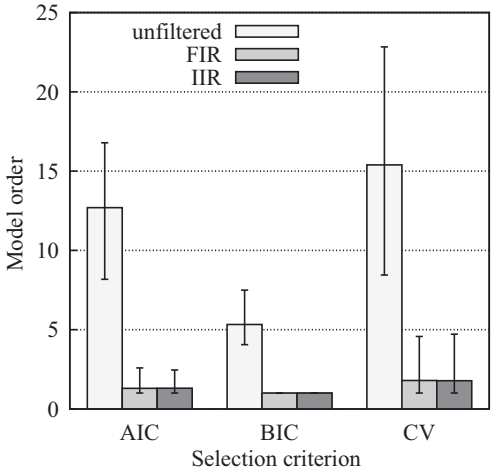

Fig. 10. Mean estimated optimal model order for unfiltered and notch-filtered time series of length 2^9, for the minimal VAR(1) with simulated line noise. Error bars indicate 95% confidence intervals.

lower model orders. The result is a very high Type I error rate (and a negligible Type II error rate). Notch filtering reduces the Type I error rate to roughly the 5% level expected for an uncontaminated signal, without increasing the Type II error rate.

4.1.4. Filter invariance and band-limited G-causality

In a final experiment, we test our proposal (Section 3.2) that an appropriate way to identify causal interactions in a specific frequency band is to use *band-limited* G-causality, i.e., in the frequency domain to simply ignore causal estimates outside the range of interest, and in the time domain to integrate over the specified range. We simulated an ensemble of 10,000 minimal VAR(1) process with significant causality $c = c_{sig}$ and other parameters as before (see introduction to Section 4.1), but with $N = 2^{12}$ time points to decrease statistical bias and improve accuracy of causal estimates. Simulations were run unfiltered and with the FIR and IIR lowpass filters as previously described (Section 4.1). Model orders were 1 for unfiltered, 44 for FIR filtered and 80 for IIR filtered data (selected as in Section 4.1.2). Spectral G-causalities were calculated with a frequency resolution of 1024 as before, and plotted (with 95% confidence intervals) against normalised frequency, along with the analytically calculated spectral G-causality of (41). We note that statistical significance testing requires permutation testing in the absence of a known null distribution for the sample estimate of band-limited G-causality (Section 2.4.1).

Results are displayed in Fig. 13. Consistent with our analysis and with Fig. 1 (Section 3), it is clear that for the (FIR and IIR) filtered data spectral G-causality is *not* suppressed in the filter stop band $\lambda \in [\pi/2, \pi]$. Note that the "ripple" on the spectral causalities for filtered data is an artefact of the in-sample finite order used to model the effectively infinite-order filtered VAR process[23] [*cf.* Section 3, in particular (26)].

Time-domain G-causalities including, for unfiltered data, band-limited causality over the pass band $\mathcal{B} = [0, \pi/2]$ calculated according to (34) by numerical quadrature, are displayed (with 95% confidence intervals) in Fig. 14. Invariance under filtering is clear within the bounds of finite-sample estimation; filtered G-causalities are close to the unfiltered G-causality value of 0.1, even though the filters strongly suppress spectral power in the filter stop band $[\pi/2, \pi]$ where spectral G-causality is much higher than in the pass band. The band-limited (unfiltered) G-causality, on the other hand, reflects correctly the much lower G-causality in the pass band. Again, accuracy of estimation—and, by implication, significance testing—is compromised by filtering. Note that the ripples evident in the spectral G-causality results (Fig. 13) do not affect the invariance in the time domain.

5. Discussion

In this paper we have shown analytically, and corroborated experimentally in simulation, that G-causality is invariant under very general digital filtering. In practice, filtering cannot and does not suppress G-causal interactions in filter stop-bands; however certain filtering operations (e.g., notch filtering) can be useful for achieving stationarity or for reducing the model order of near-non-stationary processes, facilitating better VAR model fitting and hence more accurate G-causality estimation. If the objective is to restrict G-causality inferences to specific frequency ranges, we have shown that band-limited G-causality is both theoretically valid and practicable. In the frequency domain, band-limited G-causality simply amounts to ignoring causal estimates outside the range of interest; in the time domain, it involves integrating over the range of interest.

5.1. Summary of findings and comparison with previous studies

The theoretical finding that G-causality is invariant under general filtering was alluded to almost thirty years ago (Geweke, 1982) but seems to have gone unnoticed in the subsequent literature, at least in neuroscience (Florin et al., 2010; Seth, 2010). Building on Geweke's early insight, we have shown here that the invariance arises from a generalisation of a fundamental property of G-causality, namely its invariance under the group of (unlagged) linear transformations of variables (Barrett et al., 2010). Consequently, the invariance is completely general, applying to all invertible multivariate digital filters and to time-domain, frequency-domain, and multivariate [generalised variance form (Barrett et al., 2010)] varieties of G-causality. Given this generality, how then to account for simulation results showing corruption

[23] The ripple makes filter invariance in the spectral domain less clear; in theory, at high model order (which would necessitate long time series for reliable estimation) ripples would be expected to become smaller (and higher frequency) and the invariance would be more apparent. The increased variance of the two lower plots indicates that filtering also reduces the accuracy of spectral causal estimation and thus increases the risk of mis-identification of causalities.

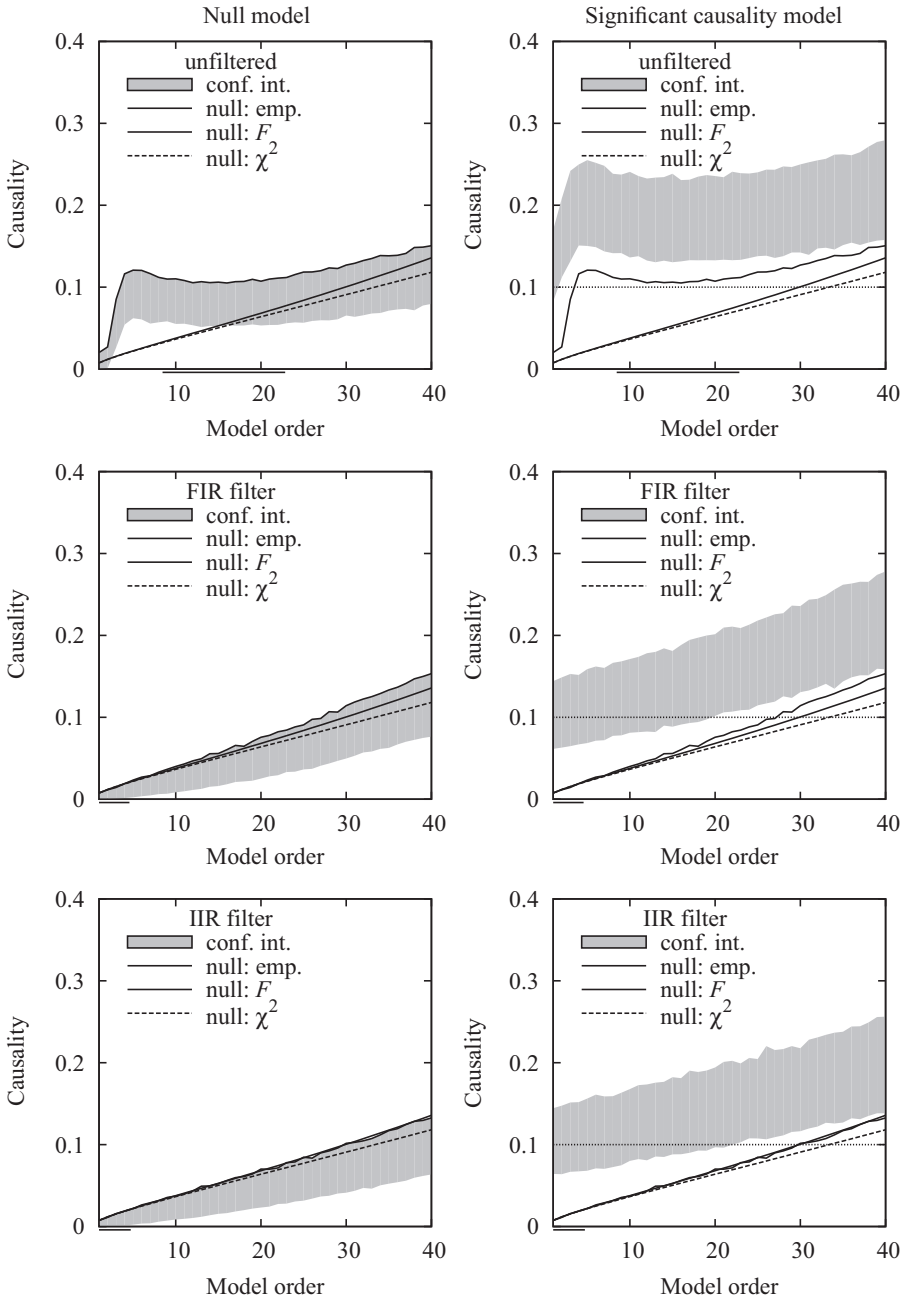

Fig. 11. Sample G-causality distributions $\mathcal{F}_{Y\to X}$ plotted against model order for filtered and unfiltered data with line noise, for the null model (left column) and significant causality model (right column). See caption to Fig. 7 for details.

of G-causal estimates following filtering (Florin et al., 2010; Seth, 2010)? We have shown that a primary cause is the large increase in empirical model induced by filtering; high model orders become necessary in order to properly fit the modified aspects of the power spectrum (low power in stop band, steep roll-off, etc.). Indeed, in theory (almost) any filtered VAR process becomes infinite order. In sample, high model orders can enable adequate model fitting but at the cost of estimating large numbers of parameters given the same data, which in turn leads to inflated Type I and Type II error rates. This explanation differs from that offered by Florin et al. (2010), who suggested that the errors are adequately accounted for by alteration of the VAR regression coefficients following filtering. Our analysis has shown that, on the theoretical level, G-causality is unaffected despite this alteration. We also identified two further potential causes of error post-filtering, (i) estimation of unstable or near-unstable models due to increased spectral radius and (ii) numerical instabilities in G-causality estimation, caused by near-vanishing spectral power. Our simulations were designed to avoid these latter two causes; however it is easy to construct examples in which they arise.

Using a minimal exactly solvable model, we investigated the effects of filtering on small-sample data with respect to statistical bias and null-hypothesis significance testing. Bias, while strong for small samples and high model orders, does not appear to be much affected by filtering. By contrast, significance testing, as reflected in Type I and Type II error rates, can be substantially affected by filtering. The inflation of error rates is apparent under the χ^2 and F-distributions, as well as under permutation sampling. Interestingly, the precise pattern of errors depends on the distribution used. For sample sizes of the order 2^9 (512) generated by our

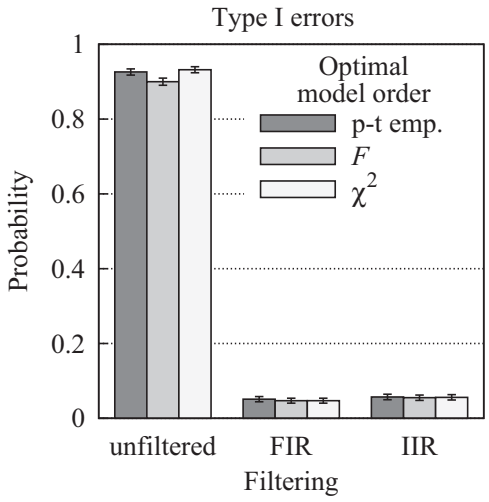

Fig. 12. Significance testing with line noise: probabilities of Type I errors (false positives) at 5% significance level for the different filter types, estimated from the permutation-test empirical ("p-t emp."), *F*- and χ^2 null distributions. Error bars indicate standard errors. Type II errors (false negatives) were insignificant in all conditions and are not shown. Model orders were the estimated optimal values: 16 for unfiltered, 2 for FIR and IIR filters.

minimal VAR(1) model, the theoretical asymptotic χ^2 distribution of the sample estimator for the G-causality statistic (under the null hypothesis of zero causality) is a poor approximation at the higher model orders implied by filtering, leading in particular to increased Type I error rates. The (non-asymptotic) *F*-distribution, while closer to the actual null distribution, also becomes less exact at higher model order leading, by contrast, to increased Type II error rates. These results imply that it is generally safer to estimate a null distribution for significance testing by a non-parametric technique such as permutation sampling. Even so, following filtering, error rates under permutation testing remain severely inflated due to the higher model orders required.

Using model orders that are substantially lower than those implied by filtering (and in our example closer to the "true" model order) had the effect of reducing the differences among the various methods of significance testing, probably reflecting a trade-off between poor estimation of a good model (high model order) and good estimation of a poor model (low model order). Error rates remain inflated in all cases, indicating that the effects of filtering on VAR modelling cannot be avoided simply by *a priori* knowledge of the underlying process.

The theoretical invariance of G-causality under filtering holds strictly for stationary processes which may be reasonably modelled as VARs, leaving open the possibility that filtering could remain useful in rendering a nonstationary sequence stationary, for example by notch filtering of (nonstationary) artifactual components such as electrical line noise, or the removal of low frequency transients by high-pass filtering; both effects might be detectable by a preliminary spectral analysis. Confirming this, we used our minimal model to show that notch filtering of data contaminated by line noise can indeed recover a stationary VAR amenable to G-causality analysis. On the other hand, we have stressed that filtering is entirely unsuitable for identifying causal interactions within specific frequency bands. Our final set of simulations shows that band-limited G-causality provides an effective alternative for estimation of time-domain G-causality within specific frequency bands.

It is worth emphasising that our simulation results were obtained using a model for which G-causality could be computed analytically. This approach allowed us to validate and explore the implications of our theoretical results with greater precision and confidence than would be possible by analysis of real data

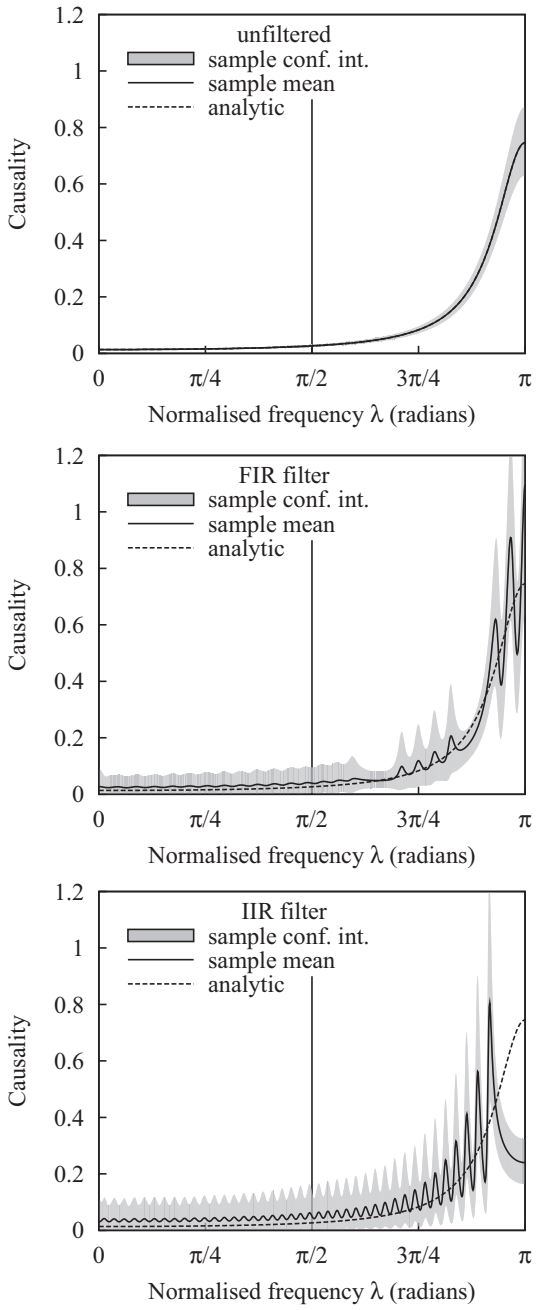

Fig. 13. Spectral G-causality for filtered and unfiltered data (solid lines) with 95% confidence intervals (shaded) estimated from a sample of 10,000 minimal VAR(1) processes, along with analytically calculated G-causality (dashed lines), plotted against normalised frequency. The vertical lines specify the lowpass filter cutoff frequency. See text for details.

or numerical simulation alone. Moreover, to our knowledge, no analytical derivation of G-causality from a generative model has previously been described in the literature. Our model therefore provides a unique platform for further methodological studies, especially on the effects of data preprocessing methods on G-causality analyses.

5.2. Related measures

In this paper we have considered only linear VAR models and G-causality measures. It might be thought that nonlinear measures (Marinazzo et al., 2011; Gourévitch et al., 2006) may offer greater robustness to filtering: since the detrimental effects of filtering fol-

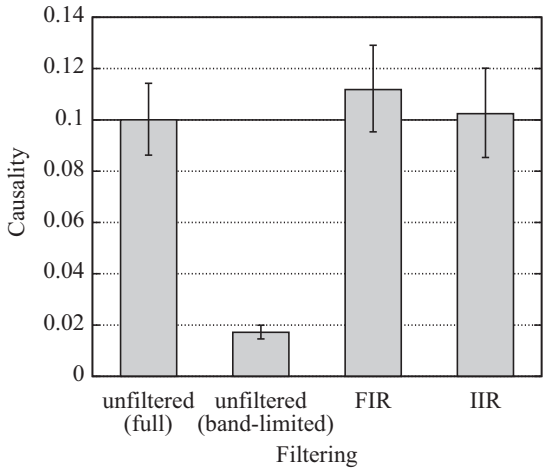

Fig. 14. Time-domain G-causalities for filtered and unfiltered data, as well as (for unfiltered data) band-limited G-causality, estimated from a sample of 10,000 minimal VAR(1) processes. Error bars indicate 95% confidence intervals. The bold horizontal line indicates the actual causality of 0.1. See text for details.

low from the large increase in empirical (linear) model order, it is plausible that nonlinear G-causality methods could lessen these effects by fitting the filter-induced power spectra distortions with fewer parameters than equivalent linear VAR models. Although we have not tested this possibility, note that the question of model fitting in sample is distinct from the question of whether linear and nonlinear measures are equivalent in theory. It is worth emphasising that for *Gaussian* processes the theoretical invariance of G-causality under filtering obtains for both linear and nonlinear versions. Indeed, we have previously shown that for Gaussian processes, G-causality and transfer entropy are equivalent, and that, furthermore, a Gaussian VAR process is necessarily linear (Barnett et al., 2009). The implication is that—at least where a Gaussian approximation is reasonable—nonlinear G-causality measures have no advantage over linear measures.

The directed transfer function (DTF) (Kaminski et al., 2001) and partially directed coherence (PDC) (Baccalá and Sameshima, 2001) are spectral causal measures related to G-causality. Preliminary investigations indicate that these measures will also be invariant, albeit under a narrower class of filtering; specifically when the component variables are independently filtered—i.e., when the filter transfer function is of the form $G(z) = \mathrm{diag}(g_1(z), \ldots, g_n(z))$ [*cf.* Eq. (25)]—and the individual filters have been normalised so that the $g_i(0)$ are all equal.

5.3. G-causality, filtering, and hemodynamic convolution

Although we have focused on filtering applied as an explicit (pre)processing step, it may be that physiological processes intervening between target neural variables and observables can impose "implicit" filtering. For example, a long-standing debate in functional connectivity analysis has been whether G-causality is appropriate for fMRI BOLD signals, given their slow dynamics relative to the underlying neural signals, the within-subject and between-subject variability in peak latency of the hemodynamic

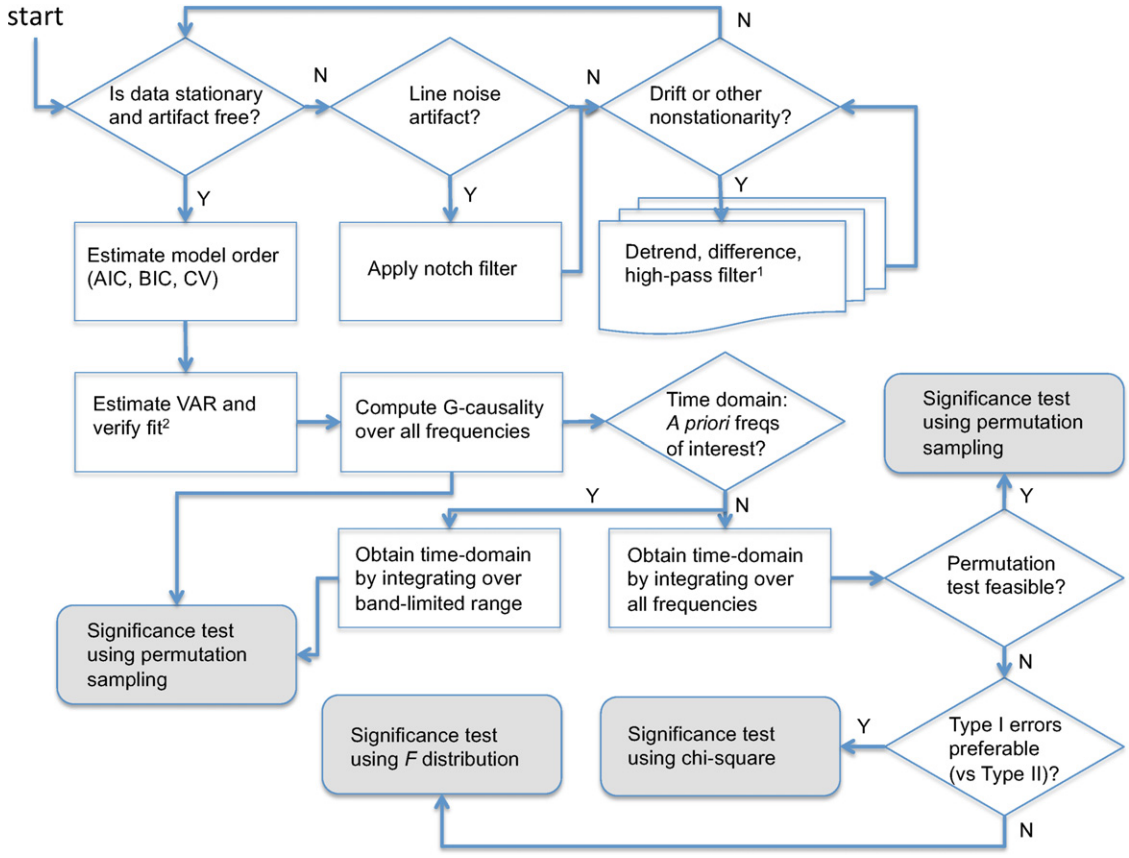

Fig. 15. Suggested pipeline for analysis. At the outset, nonstationarity and the presence of artifacts can be identified by visual inspection of the raw time series and power spectra, by VAR model stability checks as described in Section 2, and by dedicated stationarity tests such as the ADF and KPSS tests (Seth, 2010). [1]Carry out these preprocessing steps in the order suggested, each time checking if data has become stationary. [2]VAR model fit can be verified by a number of methods including R^2, model consistency, and whiteness of residuals (Seth, 2010). Recall that Type I errors are false positives, and Type II errors are false negatives (misses).

response function, and the severe downsampling imposed by scanner repetition times (David et al., 2008; Bressler and Seth, 2011; Roebroeck et al., 2011; Valdes-Sosa et al., 2011). A common suspicion is that the variation in hemodynamic latency in particular presents a fatal confound to G-causality analysis, since hemodynamic "delays" are often longer than underlying neural delays. According to this suspicion, if hemodynamic latencies oppose neural latencies, G-causality will fail to detect the "true" causality (determined by the neural delays) because it will be determined instead by the confounding hemodynamic latencies. Some simulation studies support this suspicion (Smith et al., 2011), while others show a surprising resilience of G-causality analysis under varying hemodynamics (Deshpande et al., 2010; Schippers et al., 2011). However, the hemodynamic response is often modelled as a *convolution* rather than as an explicit buffering or delay of a neural signal (Friston et al., 2000; Zumer et al., 2010), suggesting that G-causality of fMRI BOLD signals may enjoy invariance with respect to the underlying neural signals, insofar as the HRF convolution represents a stable, invertible filter. While stability is likely to hold (since, as remarked in Section 3, FIR filters are always stable) it is less clear whether the invertibility condition will be met. In particular, if the onset of the hemodynamic response is a pure delay then, considered as a filter, it will not be invertible (*cf.* Section 3). Nonetheless, this suggests a useful avenue for future work.

5.4. Summary

In summary, our results suggest that for G-causality analysis of data one should filter as little as possible, and only insofar as is necessary to render nonstationary data stationary. Fig. 15 provides a suggested pipeline for G-causality analysis, as implied by our results. Note that this flowchart has heuristic value only; it is not guaranteed to furnish valid results for all data, and for any given data set other analysis pipelines may be equally or more appropriate. Nonetheless, we hope it will provide a practically useful framework for researchers interested in inferring causal interactions from their data.

Acknowledgements

LB and AKS are supported by a donation from the Dr. Mortimer and Theresa Sackler Foundation. AKS is also supported by EPSRC fellowship EP/GOO7543/1.

References

Anderson MJ, Robinson J. Permutation tests for linear models. Aust NZ J Stat 2001;43(1):75–88.
Antoniou A. Digital filters: analysis, design, and applications. New York, NY: McGraw-Hill; 1993.
Baccalá LA, Sameshima K. Partial directed coherence: a new concept in neural structure determination. Biol Cybern 2001;84:463–74.
Barnett L, Barrett AB, Seth AK. Granger causality and transfer entropy are equivalent for Gaussian variables. Phys Rev Lett 2009;103(23):238701.
Barrett AB, Barnett L, Seth AK. Multivariate Granger causality and generalized variance. Phys Rev E 2010;81(4):41907.
Bressler S, Seth A. Wiener–Granger causality: A well established methodology. Neuroimage 2011;58(2):323–9.
Cui J, Xu L, Bressler SL, Ding M, Liang H. BSMART: a Matlab/C toolbox for analysis of multichannel neural time series. Neural Netw 2008;21:1094–104.
David O, Guillemain I, Saillet S, Reyt S, Deransart C, Segebarth C, et al. Identifying neural drivers with functional MRI: an electrophysiological validation. PLoS Biol 2008;6:2683–97.
Deshpande G, Sathian K, Hu X. Effect of hemodynamic variability on Granger causality analysis of fMRI. Neuroimage 2010;52:884–96.
Dhamala M, Rangarajan G, Ding M. Analyzing information flow in brain networks with nonparametric Granger causality. Neuroimage 2008a;41:354–62.
Dhamala M, Rangarajan G, Ding M. Estimating Granger causality from Fourier and wavelet transforms of time series data. Phys Rev Lett 2008b;100:018701.
Ding M, Bressler S, Yang W, Liang H. Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: data prepocessing, model validation, and variability assessment. Biol Cybern 2000;83:35–45.
Ding M, Chen Y, Bressler S. Granger causality: basic theory and application to neuroscience. In: Schelter S, Winterhalder M, Timmer J, editors. Handbook of time series analysis. Wienheim: Wiley; 2006. p. 438–460.
Efron B. p. 38 The jackknife, the bootstrap, and other resampling plans. In: Society of industrial and applied mathematics CBMS-NSF monographs; 1982.
Florin E, Gross J, Pfeifer J, Fink GR, Timmermann L. The effect of filtering on Granger causality based multivariate causality measures. Neuroimage 2010;50(2):577–88.
Friston KJ, Mechelli A, Turner R, Price CJ. Nonlinear responses in fMRI: the Balloon model, Volterra kernels, and other hemodynamics. Neuroimage 2000;12:466–77.
Geweke J. Measurement of linear dependence and feedback between multiple time series. J Am Stat Assoc 1982;77(378):304–13.
Geweke J. Measures of conditional linear dependence and feedback between time series. J Am Stat Assoc 1984;79(388):907–15.
Gourévitch B, Bouquin-Jeannès RL, Faucon G. Linear and nonlinear causality between signals: methods, examples and neurophysiological applications. Biol Cybern 2006;95:349–69.
Hamilton JD. Time series analysis. Princeton, NJ: Princeton University Press; 1994.
Hurvich CM, Tsai CL. Regression and time series model selection in small samples. Biometrika 1989;76:297–307.
Kaiser A, Schreiber T. Information transfer in continuous processes. Physica D 2002;166:43–62.
Kaminski M, Ding M, Truccolo WA, Bressler SL. Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. Biol Cybern 2001;85:145–57.
Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 1958;53(282):457–81.
Ladroue C, Guo S, Kendrick K, Feng J. Beyond element-wise interactions: Identifying complex interactions in biological processes. PLoS One 2009;4, e6899–e6899.
Marinazzo D, Liao W, Chen H, Stramaglia S. Nonlinear connectivity by Granger causality. Neuroimage 2011;58(2):330–8.
McQuarrie ADR, Tsai CL. Regression and time series model selection. Singapore: World Scientific Publishing; 1998.
Pollonini L, Pophale S, Situ N, Wu MH, Frye RE, Leon-Carrion J, et al. Information communication networks in severe traumatic brain injury. Brain Topogr 2010;23:221–6.
Roebroeck A, Formisano E, Goebel R. The identification of interacting networks in the brain using fMRI: model selection, causality and deconvolution. Neuroimage 2011;58(2):296–302.
Schippers MB, Renken R, Keysers C. The effect of intra- and inter-subject variability of hemodynamic responses on group level Granger causality analyses. Neuroimage 2011;57(1):22–36.
Schreiber T. Measuring information transfer. Phys Rev Lett 2000;85(2):461–4.
Seth AK. A MATLAB toolbox for Granger causal connectivity analysis. J Neurosci Methods 2010;186:262–73.
Smith SM, Miller KL, Salimi-Khorshidi G, Webster M, Beckmann CF, Nichols TE, et al. Network modelling methods for FMRI. Neuroimage 2011;54:875–91.
Valdes-Sosa PA, Roebroeck A, Daunizeau J, Friston K. Effective connectivity: influence, causality and biophysical modeling. Neuroimage 2011;58(2):339–61.
Whittle P. On the fitting of multivariate autoregressions, and the approximate canonical factorization of a spectral density matrix. Biometrika 1963;50(1–2):129–34.
Wilson DA, Yan X. Sleep-like states modulate functional connectivity in the rat olfactory system. J Neurophysiol 2010;104:3231–9.
Wilson GT. The factorization of matricial spectral densities. SIAM J Appl Math 1972;23(4):420–6.
Zumer JM, Brookes MJ, Stevenson CM, Francis ST, Morris PG. Relating BOLD fMRI and neural oscillations through convolution and optimal linear weighting. Neuroimage 2010;49:1479–89.