Reliability of the Granger causality inference

To cite this article: Douglas Zhou *et al* 2014 *New J. Phys.* **16** 043016

View the article online for updates and enhancements.

You may also like

Reliability of the Granger causality inference

Douglas Zhou[1], Yaoyu Zhang[1], Yanyang Xiao[1] and David Cai[1,2,3]

[1] Department of Mathematics, MOE-LSC, and Institute of Natural Sciences, Shanghai Jiao Tong University, Shanghai 200240, People's Republic of China
[2] Courant Institute of Mathematical Sciences & Center for Neural Science, New York University, New York, NY 10012, USA
[3] NYUAD Institute, New York University Abu Dhabi, PO Box 129188, Abu Dhabi, UAE
E-mail: zdz@sjtu.edu.cn and cai@cims.nyu.edu

Abstract

How to characterize information flows in physical, biological, and social systems remains a major theoretical challenge. Granger causality (GC) analysis has been widely used to investigate information flow through causal interactions. We address one of the central questions in GC analysis, that is, the reliability of the GC evaluation and its implications for the causal structures extracted by this analysis. Our work reveals that the manner in which a continuous dynamical process is projected or coarse-grained to a discrete process has a profound impact on the reliability of the GC inference, and different sampling may potentially yield completely opposite inferences. This inference hazard is present for both linear and nonlinear processes. We emphasize that there is a hazard of reaching incorrect conclusions about network topologies, even including statistical (such as small-world or scale-free) properties of the networks, when GC analysis is blindly applied to infer the network topology. We demonstrate this using a small-world network for which a drastic loss of small-world attributes occurs in the reconstructed network using the standard GC approach. We further show how to resolve the paradox that the GC analysis seemingly becomes less reliable when more information is incorporated using finer and finer sampling. Finally, we present strategies to overcome these inference artifacts in order to obtain a reliable GC result.

1. Introduction

Information flow plays a central role in physical and biological systems, such as in gene regulation and cortical computation in the brain. How to characterize information propagation within dynamical systems is a great theoretical challenge. There has been a long history of studying causal relations between time series [1–6]. Recently, Granger causality (GC) analysis has proven itself as a powerful method, widely used for extracting causal connectivity from data in fields ranging from physics [7–10] and biology [11–14], to social science [15–18].

In general, it is often quite involved to acquire reliable GC inference. Usually the GC framework requires the time series to be linear. For linear systems, the directed GC corresponds to the structural connectivity that mediates physical interactions within a system. For nonlinear systems, the relationship between the causal connectivity and the structural connectivity is still under active investigation. Extending the GC theory to nonlinear, non-Gaussian times series remains a central theoretical endeavor [19–21]. Note that the notion of GC is statistical rather than structural, i.e. it identifies directed statistical causal connectivity through statistical features of time series. As a statistical method, one of the important issues is how to collect time series for dynamical variables in order to faithfully capture the information flow within the dynamics. In this work, we study the reliability of GC inference for both linear and nonlinear systems and demonstrate that there are inference hazards in the GC analysis arising from the manner in which a continuous dynamical process is projected to a discrete process.

Because most dynamical quantities are continuous in time and GC values are generally evaluated using discrete time series sampled from these continuous time processes, we investigate the reliability of GC evaluation by studying the GC value as a function of the sampling interval length τ (we will term this function *the GC sampling structure*). We show that, for both linear and nonlinear dynamics, there are surprisingly common features in the GC sampling structure: (i) when by the design of our system there is a causal flow, yet the GC value can vanish over a certain set of τ, yielding an incorrect causal inference; or (ii) conversely, when there is no causal influence by construction, the GC value can become of appreciable size for some ranges of τ, yielding again a potentially erroneous causal inference. Clearly, these phenomena greatly complicate the interpretation of the GC inference and potentially produce opposing causality conclusions when using different sampling τ's for empirical data. These issues have yet to receive sufficient attention despite the wide use of GC analysis in many fields. It is important to examine these issues in a GC analysis to avoid erroneous causal inference. We use a small-world network reconstruction as an example to illustrate this hazard. As is well known, many important biological or social networks possess small-world attributes [22, 23]. The small-world topology is generally linked to efficient parallel information processing on both local and global levels. We show that an incorrect inference can arise, i.e. the small-world attributes are completely lost in the reconstructed network when GC analysis is blindly applied to the network topology reconstruction. Clearly, these sampling issues should also arise in data processing for GC evaluations using low-pass filtering or down-sampling for many types of imaging data, such as the BOLD signal in functional magnetic resonance imaging (fMRI) [24, 25].

To remove the inference artifacts in the GC analysis, it seems that a natural resolution is to incorporate as much information as possible by using sufficiently fine sampling. However, as we show below, this approach produces a paradox, i.e. a finer sampling may not imply a more reliable GC inference. In what follows, we will use idealized models to study the mechanisms underlying these phenomena and discuss an approach that can eliminate these artifacts to obtain a reliable GC relation.

2. GC sampling structures

GC aims to analyze the causal influence of one time series, X_t, on another, Y_t, by examining whether the statistical prediction of Y_t can be improved when the past information of X_t is incorporated into it [2, 26, 27]. First, to fix the notation, we briefly recapitulate the GC analysis [2, 27, 28]. For two stationary zero-mean discrete time series $\{X_t\}$ and $\{Y_t\}$, a statistical prediction can usually be accomplished using the autoregression

$$X_t = \sum_{j=1}^{\infty} a_{1j}X_{t-j} + \epsilon_{1t},$$

$$Y_t = \sum_{j=1}^{\infty} d_{1j}Y_{t-j} + \eta_{1t},$$

where ϵ_{1t} and η_{1t} are residuals representing the prediction errors when we separately consider the history of each time series. The corresponding joint regression is

$$X_t = \sum_{j=1}^{\infty} a_{2j}X_{t-j} + \sum_{j=1}^{\infty} b_{2j}Y_{t-j} + \epsilon_{2t},$$

$$Y_t = \sum_{j=1}^{\infty} c_{2j}X_{t-j} + \sum_{j=1}^{\infty} d_{2j}Y_{t-j} + \eta_{2t},$$

where ϵ_{2t} and η_{2t} are residuals representing the prediction errors after we consider the shared history for both time series. The sequences $\{\epsilon_{1t}\}$, $\{\epsilon_{2t}\}$, $\{\eta_{1t}\}$, and $\{\eta_{2t}\}$ are white-noise time series. The GC $F_{y \to x}$ characterizing the causal influence of $\{Y_t\}$ on $\{X_t\}$ is defined as

$$F_{y \to x} = \ln \frac{\text{var}(\epsilon_{1t})}{\text{var}(\epsilon_{2t})}. \tag{1}$$

Note that $\text{var}(\epsilon_{1t}) \geqslant \text{var}(\epsilon_{2t})$, i.e. the prediction of one time series cannot become worse (i.e. have greater residual variance) after information from the other time series is incorporated, therefore, $F_{y \to x} \geqslant 0$. The GC from $\{X_t\}$ to $\{Y_t\}$ is defined similarly as $F_{x \to y} = \ln \frac{\text{var}(\eta_{1t})}{\text{var}(\eta_{2t})}$. The total GC is $F_{x,y} = F_{y \to x} + F_{x \to y} + F_{x \cdot y}$, where $F_{x \cdot y}$ is the instantaneous GC, $F_{x \cdot y} = \ln \frac{\text{var}(\epsilon_{2t})\text{var}(\eta_{2t})}{|\Sigma|}$, $|\Sigma| = \text{var}(\epsilon_{2t})\text{var}(\eta_{2t}) - \left[\text{cov}(\epsilon_{2t}, \eta_{2t})\right]^2$ [2, 27, 28]. In addition, we can also describe GC in the frequency domain [2, 27, 28], e.g. $F_{x,y} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln \left(\frac{S_{yy}(\omega)S_{xx}(\omega)}{|\mathbf{S}(\omega)|} \right) d\omega$. Here, $|\mathbf{S}(\omega)|$ is the

determinant of the spectral matrix $\mathbf{S}(\omega)$, where matrix elements $S_{xx}(\omega)$ and $S_{yy}(\omega)$ are the power spectra of $\{X_t\}$ and $\{Y_t\}$, respectively, and $S_{xy}(\omega) = \sum_{n=-\infty}^{+\infty} \mathrm{cov}(X_t, Y_{t-n}) e^{-in\omega}$ is the cross-spectrum of $\{X_t\}$ and $\{Y_t\}$.

To illustrate the application of GC in nonlinear dynamics, we choose network dynamics of conductance-based, pulse-coupled integrate-and-fire (I&F) neurons [29]

$$\frac{\mathrm{d}v_i}{\mathrm{d}t} = -g_{\mathrm{L}}\left(v_i - \epsilon_{\mathrm{L}}\right) - g_i\left(v_i - \epsilon_{\mathrm{E}}\right), \tag{2a}$$

$$\frac{\mathrm{d}g_i}{\mathrm{d}t} = -\frac{g_i}{\sigma} + \sum_{j,k} s_{ji}\delta\left(t - T_{j,k}\right) + \lambda \sum_l \delta\left(t - T_{i,l}^{\mathrm{p}}\right), \tag{2b}$$

where v_i and g_i are the membrane potential and conductance for the ith neuron, respectively, g_{L} is the leak conductance, ϵ_{L} and ϵ_{E} are the resting and reversal potentials, respectively, and σ is the conductance time constant. When $v_i \in (\epsilon_{\mathrm{L}}, V_{\mathrm{th}})$, a neuron evolves according to equation (2). When v_i reaches the threshold V_{th}, it is reset to the resting potential ϵ_{L} for a refractory period τ_{ref}. This threshold-reset event is referred to as a spiking event. The coefficient $s_{ji} = sA_{ji}$ denotes the connection from neuron j to neuron i, with s the coupling strength and $\mathbf{A} = (A_{ji})$ the adjacency matrix of the network; $T_{j,k}$ is the kth spike time of the jth neuron, whereas $T_{i,l}^{\mathrm{p}}$ is the lth spike time in the Poisson input to the ith neuron. This input has rate μ and strength λ.

For a network of two neurons, x and y, with the adjacency matrix $\mathbf{A}$ with elements $A_{xx} = 0$, $A_{xy} = 0$, $A_{yx} = 1$, and $A_{yy} = 0$, we apply the GC analysis on the voltage time series. We expect that a successful GC analysis of the network topology should result in consistency between the causal and structural connectivities, that is, the presence of a structural coupling, $A_{yx} = 1$, should induce a causal flow, i.e. $F_{y \to x} \neq 0$, whereas the absence of the coupling $A_{xy} = 0$ should yield $F_{x \to y} = 0$. We demonstrate below how we can reliably assess this consistency by first addressing artifacts arising from the sampling in the evaluation of GC values.

Note that the discrete time series used in the GC analysis is usually sampled from continuous-time processes. The sampling effects can be captured by the GC sampling structure, i.e. the GC values, $F_{y \to x}$, $F_{x \to y}$, as functions of the sampling interval τ, denoted by $F_{y \to x}^{(\tau)}$, $F_{x \to y}^{(\tau)}$. For the nonlinear network dynamics of two neurons, from its GC sampling structure as shown in figure 1(a), we can observe the following properties:

 (I) GC is an oscillatory, decaying function of τ,
(II) GC becomes nearly zero, or zero, periodically.

From this GC sampling structure, clearly, one is confronted with the question of what a proper τ is to obtain the GC value for causal inference or network topology extraction.

Properties I–II are general, not limited to nonlinear dynamics, and they also appear in many of the linear models we have examined. Figure 1(b) illustrates such an example of linear (second-order regressive) models:

$$X_t = 0.9X_{t-1} - 0.6X_{t-2} - 0.3Y_{t-1} + 0.15Y_{t-2} + \varepsilon_t, \tag{3a}$$
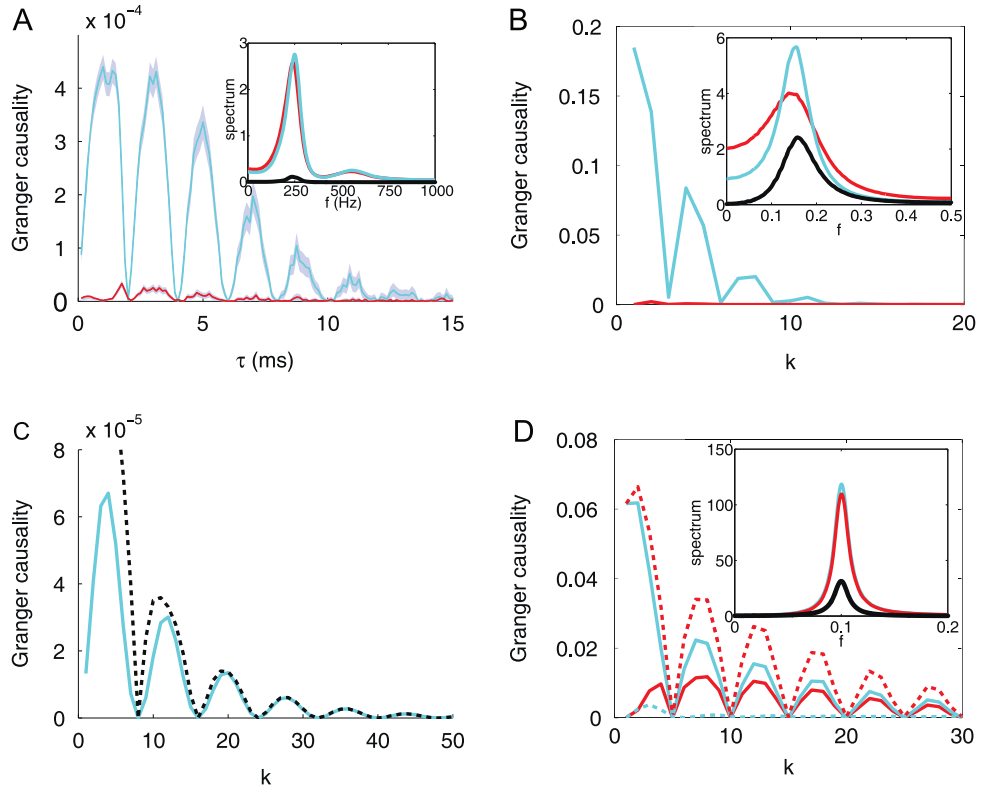
Figure 1. The GC sampling structure. (a) For the uni-directed two-neuron network (2), plotted are $F_{x \to y}^{(\tau)}$ (red) and $F_{y \to x}^{(\tau)}$ (cyan), obtained from voltage time series of the system (2). Shaded regions along the curves indicate a 95% confidence interval. The parameters are $V_{th} = 1$, $\epsilon_L = 0$, $\epsilon_E = 14/3$, $\tau_{ref} = 2$ ms, $\sigma = 2$ ms and $g_L = 0.05$ ms^{-1} throughout the text. Here $\mu = 1$ kHz, $\lambda = 0.06$, $s = 0.02$. (b) For the second-order linear regression model (3), plotted are $F_{x \to y}^{(k)}$ (red) and $F_{y \to x}^{(k)}$ (cyan) with sampling interval k. The insets in (a) and (b) are corresponding spectra: $S_{xx}(\omega)$ (cyan), $S_{yy}(\omega)$ (red), and $\left| S_{xy}(\omega) \right|$ (black) computed from the original time series. (c) Comparison with asymptotic results of $F_{y \to x}^{(k)}$: the black dotted line is numerically obtained GC and the cyan is the asymptotic form (5) with $\beta = \pi/8$ and $\phi = -\pi/2$. (d) For model II, plotted are $F_{x \to y}^{(k)}$ (red), $F_{y \to x}^{(k)}$ (cyan), $F_{x.y}^{(k)}$ (dashed cyan), $F_{x.y}^{(k)}$ (dashed red) with sampling interval k (parameters $\xi = 0.05$, $\beta = \frac{\pi}{5}$, $\gamma = 0.3$, $\phi = \frac{\pi}{2}$). Corresponding spectra are plotted in the inset: S_{xx} (cyan), S_{yy} (red), and $\left| S_{xy} \right|$ (black).

$$Y_t = 0.7 Y_{t-1} - 0.4 Y_{t-2} + \eta_t, \tag{3b}$$

where ϵ_t and η_t are Gaussian white noise with zero mean, $\mathrm{var}(\epsilon_t) = 0.5$, $\mathrm{var}(\eta_t) = 1$, and $\mathrm{cov}(\epsilon_t, \eta_t) = 0.2$. Here, 'sampling interval length' $(\tau = k)$ refers to a 'coarse-grained' discrete time series constructed using an original datum point after skipping $k - 1$ points in between. As shown in figure 1(b), clearly, the model also possesses properties I–II.

Next, we construct *idealized models* for which we have perfect control of causal influence and use them to ascertain the phenomena observed in the GC sampling structure analytically. Using the lag operator L, i.e. $LX_t = X_{t-1}$, and defining $a(L) = 1 + \sum_{n=1}^{+\infty} a_n L^n$, $b(L) = \sum_{n=1}^{+\infty} b_n L^n$, and $c(L) = 1 + \sum_{n=1}^{+\infty} c_n L^n$, the dynamics we consider are

$$\begin{aligned} X_t &= a(L)\epsilon_t + b(L)\eta_t, \\ Y_t &= c(L)\eta_t, \end{aligned} \tag{4}$$

where ϵ_t and η_t are independent white-noise series, so that $\mathrm{cov}(\epsilon_t, \eta_t) = 0$. Clearly, the structural connectivity in the system (4) means that Y_t causally influences X_t and is not influenced by X_t.

We first consider model I: $c(L) = 1$, $\mathrm{var}(\epsilon_t) = \Omega$, $\mathrm{var}(\eta_t) = 1$. Denoting the Fourier transforms of $a(L)$ and $b(L)$ by $a(\omega) = 1 + \sum_{n=1}^{+\infty} a_n e^{-in\omega}$ and $b(\omega) = \sum_{n=1}^{+\infty} b_n e^{-in\omega}$, respectively, we obtain the spectra of $\{X_t, Y_t\}$ as $S_{xx} = a(\omega)a^*(\omega)\Omega + b(\omega)b^*(\omega)$, $S_{yy} = 1$, and $S_{xy} = b(\omega)$. Here, we first consider $S_{xx} = C$ (C is a constant), i.e. there are to be no oscillations in $\{X_t\}$. Note that $S_{xx}(\omega)$ is the spectrum of the original time series $\{X_0, X_1, X_2, \cdots\}$. We denote by $S_{xx}^{(k)}(\omega)$ the spectrum for the time series $\{X_0, X_k, X_{2k}, \cdots\}$, sampled from the original $\{X_n\}$ with the sampling interval of length k. For this model, we can easily show that $F_{x,y}^{(k)} \equiv F_{x \to y}^{(k)} \equiv 0$, which are GC computed using the time series sampled with interval length k. Therefore, for the sampling interval k, we have

$$F_{y \to x}^{(k)} = F_{x,y}^{(k)} = -\frac{1}{2\pi} \int_{-\pi}^{\pi} \ln\left(1 - \frac{S_{xy}^{(k)}(\omega) S_{yx}^{(k)}(\omega)}{S_{xx}^{(k)}(\omega) S_{yy}^{(k)}(\omega)}\right) d\omega$$

by the definition of the total GC. The only spectrum component that varies with k is the cross-spectrum $S_{xy}^{(k)}$. To illustrate the behavior of the GC sampling structure, we choose $b(L) = \sum_{n=1}^{+\infty} e^{-\xi n} \cos(\beta n + \phi) L^n$, which reflects the typical oscillatory, decaying behavior of the correlations in the time series. Here ξ, β, and ϕ are constants. Under the condition that $C \gg |b(\omega)|^2$, using the spectral representation of $F_{y \to x}^{(k)}$ above, we can show that $F_{y \to x}^{(k)}$ has the asymptotic form

$$F_{y \to x}^{(k)} \approx e^{-2\xi k} \cos^2(\beta k + \phi)/C \tag{5}$$

for large k, which clearly possesses properties I–II. In particular, for suitable values of β and ϕ, the GC value vanishes on a certain set of k. Figure 1(c) shows such an example, which also confirms that the GC asymptotic solution for large k agrees well with the numerically obtained GC sampling structure from the time series. By the model construction (4), there is a causal flow from Y_t to X_t. However, through our asymptotic analysis, we confirm that the corresponding GC vanishes on certain sampling intervals, k.

To emphasize another kind of sampling artifact [30] in the GC sampling structure, we consider model II, for which $a(L) = 1 + \sum_{n=1}^{+\infty} e^{-\xi n} \cos(\beta n) L^n$, $b(L) = \gamma \sum_{n=1}^{+\infty} e^{-\xi n} \cos(\beta n + \phi) L^n$, and $c(L) = 1 + \sum_{n=1}^{+\infty} e^{-\xi n} \cos(\beta n) L^n$, where ξ, β, γ, ϕ are constants:

$\mathrm{var}\left(\epsilon_t\right) = 1$, $\mathrm{var}\left(\eta_t\right) = 1$, and $\mathrm{cov}\left(\epsilon_t, \eta_t\right) = 0$. As shown in figure 1(d), model II also possesses properties I–II. Importantly, it displays a prominent feature of spurious GC as property

(III) $F_{x\to y}^{(k)}$ (the red curve in figure 1(d)) has a non-zero value, which becomes appreciable for $k > 1$, despite the fact that there is no causal influence from $\left\{X_t\right\}$ to $\left\{Y_t\right\}$ by the model construction.

This phenomenon is also discernible in figure 1(a) for the nonlinear network, for which $A_{xy} = 0$, yet $F_{x\to y}^{(k)}$ can significantly deviate from zero over some interval τ, as clearly demonstrated by the 95% confidence region in figure 1(a).

As described above, properties I–III in the GC sampling structure are common features for both linear and nonlinear models. These properties have strong ramifications in the application of the GC analysis. Importantly, properties II and III demand additional sampling criteria for establishing the reliability of any conclusions about the causal influence for both linear and nonlinear systems.

3. Small τ limit for GC

Because GC is not invariant with respect to τ, we cannot choose the sampling-interval length arbitrarily in the GC analysis. One possible solution is to use a discrete time series sampled with very fine intervals from a continuous-time process. Intuitively, this approach would incorporate ever more information for causality determination with ever finer intervals. However, as shown in figure 2(a), for the case of nonlinear network dynamics, we have the following phenomenon as property

(IV) the GC value, obtained using a discrete time series with the sampling interval τ, approaches zero almost *linearly* as $\tau \to 0$.

Therefore, the GC constructed in this way would give rise to a paradox that we eventually lose the ability for causal inference through GC despite the fact that more information is incorporated as $\tau \to 0$. We present the mathematical reasons below for property IV, from which one can see that it is also a general phenomenon.

Because GC can be computed through the spectral representation, we first study the limit of the spectral matrix $\mathbf{S}^{(\tau)}\left(\omega\right)$ which is the spectrum of the time series $\left\{X_{n\tau}, Y_{n\tau}\right\}$, $n \in \mathbb{Z}$, with the sampling interval τ obtained from a bivariate continuous-time stationary process $\left\{X_t, Y_t\right\}$, $t \in \mathbb{R}$. Then, the covariance matrix $\mathbf{G}\left(n\tau\right)$, $n \in \mathbb{Z}$ is a sampling of

$$\mathbf{G}\left(s\right) = \begin{bmatrix} \mathrm{cov}\left(X_t, X_{t-s}\right) & \mathrm{cov}\left(X_t, Y_{t-s}\right) \\ \mathrm{cov}\left(Y_t, X_{t-s}\right) & \mathrm{cov}\left(Y_t, Y_{t-s}\right) \end{bmatrix}, \quad s \in \mathbb{R}. \quad \text{Define} \quad \mathbf{P}\left(f\right) = \int_{-\infty}^{+\infty} \mathbf{G}\left(s\right) e^{-\mathrm{i}s2\pi f} \mathrm{d}s, \quad \text{i.e. the}$$

power spectral density (PSD) of the continuous processes X_t, Y_t. By the Wiener–Khinchin theorem, $\mathbf{S}^{(\tau)}\left(\omega\right) = \sum_{n\in\mathbb{Z}} \mathbf{G}\left(n\tau\right) e^{-\mathrm{i}n\omega}$, noting that the relation between the frequency f for the continuous process and ω for the discrete time series is $\omega = 2\pi\tau f$, we obtain $\tau\mathbf{S}^{(\tau)}\left(\omega\right) \to \mathbf{P}\left(f\right)$ as $\tau \to 0$. Using this relation, the total GC spectral representation yields

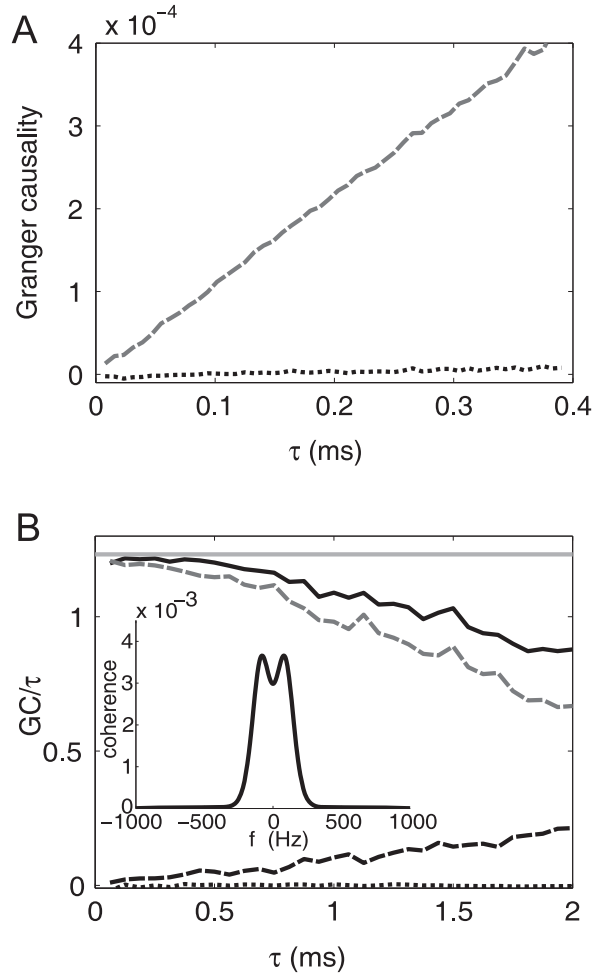**Figure 2.** The GC sampling structure as $\tau \to 0$ for the uni-directed two-neuron network (2) with $\mu = 1$ kHz, $\lambda = 0.0177$, $s = 0.02$. (a) $F_{y \to x}^{(\tau)}$ (gray dashed line) and $F_{x \to y}^{(\tau)}$ (black dotted line) obtained from voltage times series. (b) $\frac{1}{\tau}F_{x \to y}^{(\tau)}$ (dotted line), $\frac{1}{\tau}F_{y \to x}^{(\tau)}$ (gray dashed line), $\frac{1}{\tau}F_{x \cdot y}^{(\tau)}$ (black dashed line), $\frac{1}{\tau}F_{x,y}^{(\tau)}$ (black solid line), and the horizontal gray straight line is $-\int_{-\infty}^{+\infty} \ln\left(1 - C\left(f\right)\right) \mathrm{d}f$ (see equation (6)). The inset is the coherence $C(f)$ computed from PSD. Note that, by the asymptotic distribution theory of GC [27], the estimator of a directed GC has a bias p/n, where p is the regression order and n is the length of the time series. We have used the Bayesian information criterion [31] to determine the regression order p and have subtracted this type of bias in (a) and (b).

$$\frac{1}{\tau}F_{x,y}^{(\tau)} \to - \int_{-\infty}^{+\infty} \ln\left(1 - C\left(f\right)\right)\mathrm{d}f \text{ as } \tau \to 0 \tag{6}$$

where $C\left(f\right) = \left[P_{xy}\left(f\right)P_{yx}\left(f\right)\right] / \left[P_{xx}\left(f\right)P_{yy}\left(f\right)\right]$ is the coherence of the continuous-time processes X_t and Y_t. Similarly, we can also show that $\frac{1}{\tau}F_{x \to y}^{(\tau)}$, $\frac{1}{\tau}F_{y \to x}^{(\tau)}$, and $\frac{1}{\tau}F_{x \cdot y}^{(\tau)}$ have limits as $\tau \to 0$ and these limits are intrinsic properties of the continuous processes. Therefore, the GC is linearly proportional to the sampling interval length for small τ.

4. Procedure for reliable GC inference

From the above analysis, it is clear that one needs to be rather careful in interpreting causal inference using the GC analysis. Our analysis also provides a general approach to circumventing the issues discussed above: first, a range of τ should be used to sample continuous-time processes to obtain a set of discrete time series. Second, one computes the GC structure to examine its general behavior in order to (i) avoid using an accidental sampling interval τ for which GC may vanish despite the presence of causal influence, or (ii) avoid spurious non-zero GC values for finite τ as seen in figure 1(d) when there is no causal influence. Third, on account of the typical length scales τ_{osci} and τ_{d} of the oscillations and decay in the GC sampling structure, one should use fine sampling $\tau \ll \tau_{\mathrm{osci}}, \tau_{\mathrm{d}}$ to find the linear range of the GC, $F^{(\tau)}$ (with the estimator bias [27] removed as in figure 2), as a function of τ, and then plot the ratio of $F^{(\tau)}/\tau$ to extract its limiting value as $\tau \to 0$.

Returning to the network of two neurons, using the above approach, we can now successfully confirm the consistency between the causal connectivity and structural connectivity, i.e. that the topology $A_{yx} = 1$ and $A_{xy} = 0$ of the two-neuron networks can be reliably linked to $\frac{1}{\tau}F_{y\to x}^{(\tau)} \neq 0$, $\frac{1}{\tau}F_{x\to y}^{(\tau)} = 0$ as $\tau \to 0$, as shown in figure 2(b).

In addition, we comment on some other aspects of the GC oscillation features observed in figure 1. As shown in figures 1(a) and (b) for both linear and nonlinear dynamics, the oscillation frequency in the GC sampling structure is twice the peak frequencies of the spectra. This relation also appears in the asymptotic solution (5) for model I, which contains the oscillation frequency $f_{\mathrm{GC}} = \beta/\pi$ and whose $\left|S_{xy}(\omega)\right|$ has a peak at $f_{\mathrm{S}} = \beta/2\pi = f_{\mathrm{GC}}/2$. Note that for nonlinear networks, the GC analysis extracts an underlying regression structure. The spectra (figure 1(a)) in the network case all have the same peak frequency, resembling those in model II (figure 1(d)), therefore, they share similar GC sampling structures by the GC spectral theory. Complicated oscillation structures in the GC sampling structures may arise and, in general, are related to the spectral peak frequencies of the original time series.

5. An example: small-world networks

Finally, we illustrate the importance of the above GC extraction strategy in an example of the reconstruction of a directed, small-world network of neurons. The small world network is initially wired using a similar method to the Watts–Strogatz method [22] with a modification to take into account directed edges [32, 33]. The small-world attributes are reflected in a large clustering coefficient and a small shortest path length in comparison with random networks. For our network of 100 excitatory neurons, the average clustering coefficient and the average shortest path lengths are 0.434 and 6.53, respectively. (For an undirected random network of the same number of nodes and edges, the average clustering coefficient and the average shortest path length would be 0.040, 2.44, respectively. For a directed random network, a node can then no longer always reach any other node through a directed path.) Using the voltage time series obtained from solving system (2), we can compute the conditional GC from neuron i to neuron j given the information about other neurons in the network (for details see [34]). The network topology is then successfully reconstructed (i.e. identical to that of the original network), as shown in figure 3(a), using the above GC extraction strategy. If a fixed sampling interval, say,
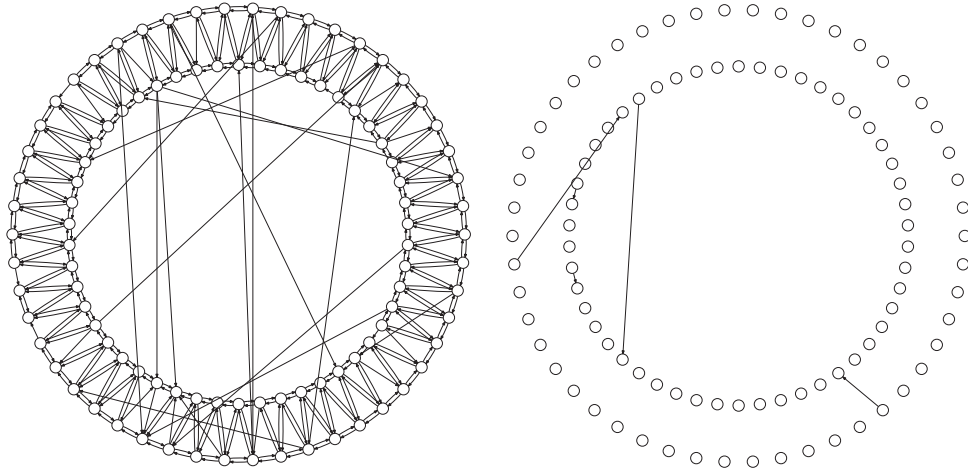
Figure 3. Loss of small-world attributes. Left panel: a small-world I&F network of 100 excitatory neurons whose directed connections are constructed using our GC extraction strategy with the smallest sampling interval $\tau = 1$ ms (see text). These have been verified to be identical to the original directed network topology; here, a neuron is represented by a small circle and a directed edge is indicated by an arrow pointing from the presynaptic neuron to the postsynaptic neuron. Right panel: a network topology constructed using the $\tau = 6$ ms sampling interval in the conventional GC analysis. This has recovered only five directed connections, while failing to capture nearly ~ 400 original connections. This drastic loss of small-world attributes is a consequence of the incorrect sampling interval used in the GC analysis. The parameters for the I&F system (2) are $\mu = 1$ kHz, $\lambda = 0.012$ and $s = 0.005$.

$\tau = 6.0$ ms, was used, the conventional GC analysis would produce a network topology that completely fails to capture the original network topology. This reconstructed network is shown in figure 3(b), which recovers only five directed connections while failing to find nearly 400 original connections in the total 9900 possible edges. Clearly, this reconstructed network, with a vanishing clustering coefficient and no connected paths, has even failed to capture the statistical structures of the network. Incidentally, we can also successfully reconstruct networks of coupled excitatory and inhibitory neurons with an arbitrary adjacent coupling matrix [19]. This example suggests that one should be very careful about the interpretation and implication of causal connectivity obtained using conventional methods for time series with a fixed sampling frequency. A systematic verification is needed to establish their validity.

6. Discussion and conclusion

In summary, we have shown that for both linear and nonlinear processes the computed GC is dependent on the sampling interval length τ. For instance, the GC may vanish on a certain set of τ despite the presence of true causal influence, or it can become non-zero over a range of τ despite the absence of any causal influence. Furthermore, the naive idea of using a sufficiently fine sampling interval, thus incorporating as much information as possible, will always give rise to vanishing GC regardless of whether there is a causal influence between the time series or not. By constructing simple idealized models, we have ascertained analytically the above phenomena as observed in both linear and nonlinear processes, and proposed approaches to

overcome these sampling artifacts in order to obtain a reliable GC inference. We discussed the sampling issues of GC using the parametric methods of GC evaluations in the time domain as above. However, we point out that, naturally, similar issues about the sampling effects in the GC analysis also exist for non-parametric methods in the frequency domain. One has to deal carefully with the sampling issue regardless of whether a parametric method or non-parametric method of computing GC is used. We further comment that there is another related, interesting scientific question, namely, even for a fixed sampling rate, whether there is causal influence from one time scale to another in a coarse-grained sense. We note that this question was addressed in [35] by using a wavelet decomposition in time. Finally, we expect that our approach of obtaining reliable GC values can be used in analyzing GC relations for various kinds of observational data, such as EEG and fMRI signals, and may shed light on the impact of sampling effects in other empirical data-inference methods.

Acknowledgments

References

[1] Orcutt G H 1952 *Rev. Econ. Stat.* **34** 305
[2] Granger C W J 1969 Investigating causal relations by econometric models and cross spectral methods *Econometrica* **37** 424–38
[3] Geweke J 1979 *Handbook of Econometrics* vol 2, 1st edn (Amsterdam: Elsevier) p 1101
[4] Schreiber T 2000 Measuring information transfer *Phys. Rev. Lett.* **85** 461–4
[5] Hlavavckova-Schindler K, Palus M, Vejmelka M and Bhattacharya J 2007 Causality detection based on information-theoretic approaches in time series analysis *Phys. Rep.* **441** 1–46
[6] Nolte G, Ziehe A, Nikulin V V, Schlogl A, Kramer N, Brismar T and Muller K R 2008 Robustly estimating the flow direction of information in complex physical systems *Phys. Rev. Lett.* **100** 234101
[7] Kaufmann R K and Stern D I 1997 Evidence for human influence on climate from hemispheric temperature relations *Nature* **388** 39–44
[8] Vanhellemont F, Didier F and Christine B 2002 Cosmic rays and stratospheric aerosols: evidence for a connection *Geophys. Res. Lett.* **29** 1715–8
[9] Ganapathy R, Rangarajan G and Sood A 2007 Granger causality and cross recurrence plots in rheochaos *Phys. Rev.* E **75** 016211
[10] Angelini L, Pellicoro M and Stramaglia S 2009 Granger causality for circular variables *Phys. Lett.* A **373** 2467–70
[11] Dhamala M, Rangarajan G and Ding M 2008 Analyzing information flow in brain networks with nonparametric Granger causality *Neuroimage* **41** 354–62
[12] Gow D W J, Segawa J A, Ahlfors S P and Lin F H 2008 Lexical influences on speech perception: a Granger causality analysis of MEG and EEG source estimates *Neuroimage* **43** 614–23

[13] Hamilton J P, Chen G, Thomason M E, Schwartz M E and Gotlib I H 2011 Investigating neural primacy in major depressive disorder multivariate Granger causality analysis of resting state fMRI time-series data *Mol. Psychiatry* **16** 763–72

[14] Hempel S, Koseska A, Kurths J and Nikoloski Z 2011 Inner composition alignment for inferring directed networks from short time series *Phys. Rev. Lett.* **107** 054101

[15] Freeman J R 1983 Granger causality and the times series analysis of political relationships *Am. J. Polit. Sci.* **27** 327–58

[16] Narayan P K and Smyth R 2004 Crime rates male youth unemployment and real income in Australia: evidence from Granger causality tests *Appl. Econ.* **36** 2079–95

[17] Tsen W H 2006 Granger causality tests among openness to international trade, human capital accumulation and economic growth in China *Int. Econ. J.* **20** 285–302

[18] Francis B M, Moseley L and Iyare S O 2007 Energy consumption and projected growth in selected caribbean countries *Energ. Econ.* **29** 1224–32

[19] Zhou D, Xiao Y, Zhang Y, Xu Z and Cai D 2013 Causal and structural connectivity of pulse coupled nonlinear networks *Phys. Rev. Lett.* **111** 054102

[20] Marinazzo D, Pellicoro M and Stramaglia S 2008 Kernel method for nonlinear Granger causality *Phys. Rev. Lett.* **100** 144103

[21] Liao W, Marinazzo D, Pan Z, Gong Q and Chen H 2009 Kernel Granger causality mapping effective connectivity on fMRI data *IEEE Trans. Med. Imaging* **28** 1825–35

[22] Watts D J and Strogatz S H 1998 Collective dynamics of small-world networks *Nature* **393** 440–2

[23] Bassett D S and Bullmore E 2006 Small-world brain network *Neuroscientist* **12** 512–23

[24] Roebroeck A, Formisano E and Goebel R 2005 Mapping directed influence over the brain using Granger causality and fMRI *Neuroimage* **25** 230–42

[25] Wen X, Rangarajan G and Ding M 2013 Is Granger causality a viable technique for analyzing fMRI data *PLoS ONE* **8** e67428

[26] Wiener N 1956 *Modern Mathematics for Engineers* (New York: McGraw-Hill)

[27] Geweke J 1982 Measurement of linear dependence and feedback between multiple time series *J. Am. Stat. Assoc.* **77** 304–13

[28] Ding M, Chen Y and Bressler S L 2006 Granger causality: basic theory and application to neuroscience *Handbook of Time Series Analysis* (Weinheim: Wiley) pp 437–60

[29] Koch C 1999 *Biophysics of Computation: Information Processing in Single Neurons* (New York: Oxford University Press)

[30] McCrorie J R and Chambers M J 2006 Granger causality and the sampling of economic processes *J. Econometrics* **132** 311–36

[31] Schwarz G 1978 Estimating the dimension of a model *Ann. Stat.* **6** 461–4

[32] Fagiolo G 2007 Clustering in complex directed networks *Phys. Rev.* E **76** 026107

[33] Liao W, Ding J, Marinazzo D, Xu Q, Wang Z, Yuan C, Zhang Z, Lu G and Chen H 2011 Small-world directed networks in the human brain: multivariate Granger causality analysis of resting-state fMRI *Neuroimage* **54** 2683–94

[34] Zhou D, Xiao Y, Zhang Y, Xu Z and Cai D 2014 Granger causality network reconstruction of conductance based integrate-and-fire neuronal systems *PLoS ONE* **9** e87636

[35] Lungarella M, Pitti A and Kuniyoshi Y 2007 Information transfer at multiple scales *Phys. Rev.* E **76** 056117