

Granger causality in the frequency domain: derivation and applications

Vinicius Lima^{*1}, Fernanda Jaiara Dellajustina¹, Renan O. Shimoura¹, Mauricio Girardi-Schappo¹, Nilton L. Kamiji¹, Rodrigo F. O. Pena¹, Antonio C. Roque¹

¹Universidade de São Paulo, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Laboratório de Sistemas Neurais, Departamento de Física, Ribeirão Preto, SP, Brasil.

Received on January 03, 2020. Revised on June 25, 2020. Accepted on July 27, 2020.

Physicists are starting to work in areas where noisy signal analysis is required. In these fields, such as Economics, Neuroscience, and Physics, the notion of causality should be interpreted as a statistical measure. We introduce to the lay reader the Granger causality between two time series and illustrate ways of calculating it: a signal X “Granger-causes” a signal Y if the observation of the past of X increases the predictability of the future of Y when compared to the same prediction done with the past of Y alone. In other words, for Granger causality between two quantities it suffices that information extracted from the past of one of them improves the forecast of the future of the other, even in the absence of any physical mechanism of interaction. We present derivations of the Granger causality measure in the time and frequency domains and give numerical examples using a non-parametric estimation method in the frequency domain. Parametric methods are addressed in the Appendix. We discuss the limitations and applications of this method and other alternatives to measure causality.

Keywords: Granger causality, autoregressive process, conditional Granger causality, non-parametric estimation

1. Introduction

The notion of causality has been the concern of thinkers at least since the ancient Greeks [1]. More recently, Clive Granger [2], in his paper entitled “Investigating Causal Relations by Econometric Models and Cross-spectral Methods” from 1969, elaborated a mathematical framework to describe a form of causality – henceforth called *Granger Causality*¹ (GC) in order to distinguish it from other definitions of causality. Given two stochastic variables, $X(t)$ and $Y(t)$, there is a causal relationship (in the sense of Granger) between them if the past observations of Y help to predict the current state of X , and *vice-versa*. If so, then we say that Y Granger-causes X . Granger was inspired by the definition of causality from Norbert Wiener [3], in which Y causes X if knowing the past of Y increases the efficacy of the prediction of the current state of $X(t)$ when compared to the prediction of $X(t)$ by the past values of X alone².

In the multidisciplinary science era, more and more physicists are involved in research in other areas, such as Economics and Neuroscience. These areas usually have big data sets. Data analysis tools, such as GC, come in handy to extract meaningful knowledge from these sets.

Causality inference via GC has been widely applied in different areas of science, such as: prediction of financial time series [7–9], earth systems [10], atmospheric systems [11], solar indices [12], turbulence [12, 13], inference of information flow in the brain of different animals [14–18], and inference of functional networks of the brain using fMRI [19, 20], MEG [21], and EEG [22]. It appears as an alternative to measures like linear correlations [23], mutual information [24, 25], partial directed coherence [26], ordinary coherence [27], directed transfer function [28], spectral coherence [29], and transfer entropy [30, 31], being usually easier to calculate since it does not rely on the estimation of probability density functions of one or more variables.

The definition of GC involves the prediction of future values of stochastic time series (see Fig. 1). The measurement of the GC between variables may be done in both the time and the frequency domains [26, 32–34].

In the present work, we will focus on the frequency domain representation of the GC [26, 32] and, for pedagogical purposes, will discuss illustrative examples from previous works by other authors [26, 34]. Our main goal is to provide a basic notion of the GC measure to a reader not yet introduced to this subject.

This work is organized as follows: in Section 2, we present the concept of an autoregressive process – a model of linear regression in which GC is based (it is also possible to formulate GC for nonlinear systems, however such a formulation results in a more complex analysis

*Correspondence email address: vinicius.lima.cordeiro@usp.br

¹It is also referred as Wiener-Granger causality.

²Other notions of causality have been defined, one worth mentioning is Pearl’s causality [4]. Over the years, Pearl’s causality has been revised by him and colleagues in a series of published works [5, 6].

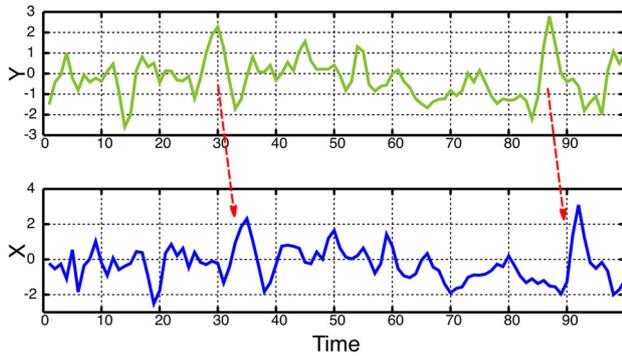


Figure 1: When time series $Y(t)$ Granger-causes time series $X(t)$, the patterns in $Y(t)$ are approximately repeated in $X(t)$ after some time lag (two examples are indicated with arrows). Thus, past values of X can be used for the prediction of future values of Y .

which is beyond the scope of this work [35,36]). Sections 3 and 4 are used to develop the mathematical concepts and definitions of the GC both in the time and frequency domains. In Section 5.1, we introduce the nonparametric method to estimate GC through Fourier and wavelet transforms [34]. In Section 6 we introduce examples of the conditional GC (cGC) to determine known links between the elements of a simple network. We then close the paper by discussing applications, implications and limitations of the method.

2. Autoregressive process

Autoregressive processes form the basis for the parametric estimation of the GC, so in this section we introduce the reader to the basic concepts of such processes [37]. A process $X(t)$ is autoregressive of order n (i.e., $AR(n)$) if its state at time t is a function of its n past states:

$$X(t) = \sum_{i=1}^n a_i X(t-i) + \epsilon(t), \tag{1}$$

where t is the integer time step, and the real coefficients a_i indicate the weighted contribution from i steps in the past, to the current state t of X . The term $\epsilon(t)$ is a noise source with variance Σ that models any external additive contribution to the determination of $X(t)$. If Σ is large, then the process is weakly dependent on its past states and $X(t)$ may be regarded as just noise. Fig. 2 shows examples of an $AR(2)$ (a) and an $AR(4)$ (b).

Fitting the autoregressive coefficients a_i and the noise variance Σ , for a recorded signal, is usually done by solving a Yule-Walker set of equations [15,38]. For a brief review on this topic see the Section A of the Appendix.

3. Granger causality in time domain

In this section we develop the mathematical concepts and definitions of GC in time domain. Consider two stochastic

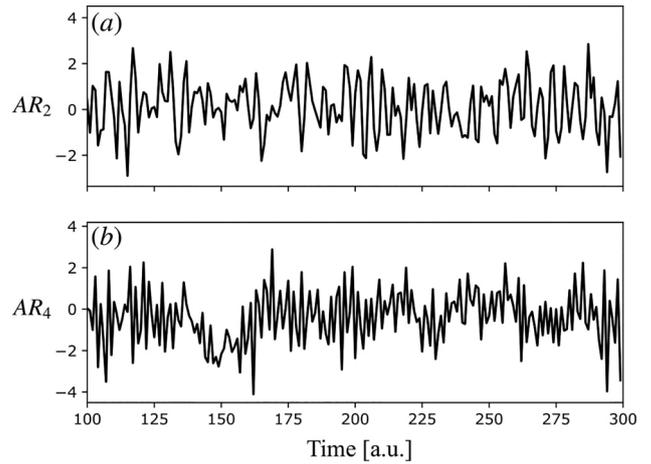


Figure 2: Autoregressive processes. (a) time series of an AR_2 process with coefficients $(a_1, a_2) = (0.3, -0.5)$. (b) time series of an AR_4 process with coefficients $(a_1, a_2, a_3, a_4) = (-0.2, 0.5, 0.6, -0.2)$.

signals, $X_1(t)$ and $X_2(t)$. We assume that these signals may be modeled by autoregressive stochastic processes of order n , independent of each other, such that their states in time t could be estimated by their n past values:

$$X_1(t) = \sum_{i=1}^n a_i X_1(t-i) + \epsilon_1(t), \tag{2}$$

$$X_2(t) = \sum_{i=1}^n c_i X_2(t-i) + \epsilon_2(t), \tag{3}$$

where the variances of ϵ_1 and ϵ_2 are, respectively, Σ_{11} and Σ_{22} , and the coefficients a_i and c_i are adjusted in order to minimize Σ_{11} and Σ_{22} .

However, we may also assume that the signals $X_1(t)$ and $X_2(t)$ are each modeled by a combination of one another, yielding

$$X_1(t) = \sum_{i=1}^n a_i X_1(t-i) + \sum_{i=1}^n b_i X_2(t-i) + \epsilon_1^*(t), \tag{4}$$

$$X_2(t) = \sum_{i=1}^n c_i X_2(t-i) + \sum_{i=1}^n d_i X_1(t-i) + \epsilon_2^*(t), \tag{5}$$

where the covariance matrix is given by

$$\Sigma = \begin{bmatrix} \Sigma_{11}^* & \Sigma_{12}^* \\ \Sigma_{21}^* & \Sigma_{22}^* \end{bmatrix}. \tag{6}$$

Here, Σ_{11}^* , Σ_{22}^* are the variances of ϵ_1^* and ϵ_2^* respectively, and $\Sigma_{12}^* = \Sigma_{21}^*$ is the covariance of ϵ_1^* and ϵ_2^* . Again, the coefficients a_i , b_i , c_i and d_i are adjusted to minimize the variances Σ_{11}^* and Σ_{22}^* .

If $\Sigma_{11}^* < \Sigma_{11}$, then the addition of $X_2(t)$ to $X_1(t)$ generated a better fit to $X_1(t)$, and thus enhanced its predictability. In this sense, we may say there is a causal relation from X_2 to X_1 , or simply that X_2 Granger-causes X_1 . The same applies for the other signal: if $\Sigma_{22}^* < \Sigma_{22}$,

then X_1 Granger-causes X_2 because adding X_1 to the dynamics of X_2 enhanced its predictability.

We may summarize this concept into the definition of the *total causality index*, given by

$$F_{1.2} = \log \left(\frac{\Sigma_{11}\Sigma_{22}}{\det(\Sigma)} \right) = \log \left(\frac{\Sigma_{11}\Sigma_{22}}{\Sigma_{11}^*\Sigma_{22}^* - (\Sigma_{12}^*)^2} \right). \quad (7)$$

If $F_{1.2} > 0$, there is some Granger-causal relation between X_1 and X_2 , because either $\Sigma_{11}^* < \Sigma_{11}$ or $\Sigma_{22}^* < \Sigma_{22}$, otherwise there is correlation between X_1 and X_2 due to $\Sigma_{12}^* > 0$. If neither Granger-causality nor correlations are present, then $F_{1.2} = 0$.

To know specifically whether there is Granger causality from 1 to 2 or from 2 to 1, we may use the specific indices:

$$F_{1 \rightarrow 2} = \log \left(\frac{\Sigma_{22}}{\Sigma_{22}^*} \right), \quad (8)$$

$$F_{2 \rightarrow 1} = \log \left(\frac{\Sigma_{11}}{\Sigma_{11}^*} \right), \quad (9)$$

$$F_{1 \leftrightarrow 2} = \log \left(\frac{\Sigma_{11}^*\Sigma_{22}^*}{\det(\Sigma)} \right), \quad (10)$$

such that

$$F_{1.2} = F_{1 \rightarrow 2} + F_{2 \rightarrow 1} + F_{1 \leftrightarrow 2}, \quad (11)$$

where $F_{1 \rightarrow 2}$ defines the causality from $X_1(t)$ to $X_2(t)$, $F_{2 \rightarrow 1}$ is the causality from $X_2(t)$ to $X_1(t)$, and $F_{1 \leftrightarrow 2}$ is called *instantaneous causality* due to correlations between ϵ_1^* and ϵ_2^* . Just as for the total causality case, these specific indices are greater than zero if there is Granger causality, or zero otherwise.

4. Granger causality in frequency domain

In order to derive the GC in frequency domain, we first define the lag operator L^k , such that

$$L^k X(t) = X(t - k), \quad (12)$$

delays $X(t)$ by k time steps, yielding $X(t - k)$. We may then rewrite equations (4) and (5) as:

$$X_1(t) = \left(\sum_{i=1}^n a_i L^i \right) X_1(t) + \left(\sum_{i=1}^n b_i L^i \right) X_2(t) + \epsilon_1^*(t), \quad (13)$$

$$X_2(t) = \left(\sum_{i=1}^n c_i L^i \right) X_1(t) + \left(\sum_{i=1}^n d_i L^i \right) X_2(t) + \epsilon_2^*(t), \quad (14)$$

and rearrange their terms to collect $X_1(t)$ and $X_2(t)$:

$$\left(1 - \sum_{i=1}^n a_i L^i \right) X_1(t) + \left(- \sum_{i=1}^n b_i L^i \right) X_2(t) = \epsilon_1^*(t), \quad (15)$$

$$\left(- \sum_{i=1}^n c_i L^i \right) X_1(t) + \left(1 - \sum_{i=1}^n d_i L^i \right) X_2(t) = \epsilon_2^*(t). \quad (16)$$

We define the coefficients $a(L) = 1 - \sum_{i=1}^n a_i L^i$, $b(L) = - \sum_{i=1}^n b_i L^i$, $c(L) = - \sum_{i=1}^n c_i L^i$ and $d(L) = 1 - \sum_{i=1}^n d_i L^i$, and rewrite equations (15) and (16) into matrix form:

$$\begin{pmatrix} a(L) & b(L) \\ c(L) & d(L) \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} = \begin{pmatrix} \epsilon_1^*(t) \\ \epsilon_2^*(t) \end{pmatrix} \quad (17)$$

where $a(0) = d(0) = 1$ and $b(0) = c(0) = 0$.

We apply the Fourier transform to equation (17) in order to switch to the frequency domain,

$$\underbrace{\begin{pmatrix} \tilde{a}(\omega) & \tilde{b}(\omega) \\ \tilde{c}(\omega) & \tilde{d}(\omega) \end{pmatrix}}_{\mathbf{A}(\omega)} \underbrace{\begin{pmatrix} X_1(\omega) \\ X_2(\omega) \end{pmatrix}}_{\mathbf{X}(\omega)} = \underbrace{\begin{pmatrix} \epsilon_1^*(\omega) \\ \epsilon_2^*(\omega) \end{pmatrix}}_{\Sigma(\omega)}, \quad (18)$$

where ω is the frequency and $\mathbf{A}(\omega)$ is the coefficient matrix whose elements are given by

$$\begin{aligned} \tilde{a}(\omega) &= 1 - \sum_{i=1}^n a_i \exp(-j\omega i), \\ \tilde{b}(\omega) &= - \sum_{i=1}^n b_i \exp(-j\omega i), \\ \tilde{c}(\omega) &= - \sum_{i=1}^n c_i \exp(-j\omega i), \\ \tilde{d}(\omega) &= 1 - \sum_{i=1}^n d_i \exp(-j\omega i). \end{aligned}$$

The expressions above are obtained by representing the lag operator in the spectral domain as $L^i = \exp(-j\omega i)$. This derives from the z -transform, where the representation of the z variable³ in the unit circle ($|z| = 1$) is $z^{-i} = \exp(-j\omega i)$ [40, 41].

To obtain the power spectra of $X_1(\omega)$ and $X_2(\omega)$, we first isolate $\mathbf{X}(\omega)$ in equation (18):

$$\begin{pmatrix} X_1(\omega) \\ X_2(\omega) \end{pmatrix} = \underbrace{\begin{pmatrix} H_{11}(\omega) & H_{12}(\omega) \\ H_{21}(\omega) & H_{22}(\omega) \end{pmatrix}}_{\mathbf{H}(\omega)} \begin{pmatrix} \epsilon_1^*(\omega) \\ \epsilon_2^*(\omega) \end{pmatrix}, \quad (19)$$

where $\mathbf{H}(\omega) = \mathbf{A}^{-1}(\omega)$ is called the transfer matrix, resulting in the following spectra:

$$\mathbf{S}(\omega) = \langle \mathbf{X}(\omega) \mathbf{X}^\dagger(\omega) \rangle = \mathbf{H}(\omega) \Sigma(\omega) \mathbf{H}^\dagger(\omega), \quad (20)$$

³The lag operator L is similar to the z -transform. However, z is treated as a variable, and is often used in signal processing, while L is an operator [39].

where $\langle \cdot \rangle$ is the ensemble average, \dagger the transposed conjugate of the matrix, and $\mathbf{S}(\omega)$ is the spectral matrix defined as:

$$\mathbf{S}(\omega) = \begin{bmatrix} S_{11}(\omega) & S_{12}(\omega) \\ S_{21}(\omega) & S_{22}(\omega) \end{bmatrix}. \quad (21)$$

In equation (21), $S_{11}(\omega)$ and $S_{22}(\omega)$ are called the autospectra, and the elements $S_{12}(\omega)$ and $S_{21}(\omega)$ are called the cross-spectra.

We can expand the product in equation (20) to obtain $S_{11}(\omega)$ and $S_{22}(\omega)$ (see Section B of the Appendix for details) as:

$$S_{11}(\omega) = \underbrace{\bar{H}_{11}(\omega)\Sigma_{11}\bar{H}_{11}^\dagger(\omega)}_{\text{Intrinsic}} + \underbrace{H_{12}(\omega)\left(\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}\right)H_{12}^*(\omega)}_{\text{Causal}}, \quad (22)$$

$$S_{22}(\omega) = \underbrace{\hat{H}_{22}(\omega)\Sigma_{22}\hat{H}_{22}^\dagger(\omega)}_{\text{Intrinsic}} + \underbrace{\bar{H}_{21}(\omega)\left(\Sigma_{11} - \frac{\Sigma_{21}^2}{\Sigma_{22}}\right)\bar{H}_{21}^*(\omega)}_{\text{Causal}}, \quad (23)$$

where the symbols $\bar{\cdot}$ and $\hat{\cdot}$ are used to differentiate the terms below from the variables H_{11} , H_{21} , and H_{22} , as follows:

$$\begin{aligned} \bar{H}_{11}(\omega) &= H_{11}(\omega) + \Sigma_{12}H_{12}(\omega)\Sigma_{11}, \\ \bar{H}_{21}(\omega) &= H_{21}(\omega) + \Sigma_{12}H_{11}(\omega)\Sigma_{11}, \\ \hat{H}_{22}(\omega) &= H_{22}(\omega) + \frac{\Sigma_{12}}{\Sigma_{22}}H_{21}(\omega). \end{aligned}$$

Once we have the $S_{11}(\omega)$ and $S_{22}(\omega)$ spectra as the sum of an *intrinsic* and a *causal* term, we may define indices to quantify GC in frequency domain just as we did in the time domain (Section 3). For instance, to calculate the causal index, we divide the spectra by their respective intrinsic term in order to eliminate its influence. Thus, the causality index $I_{2 \rightarrow 1}(\omega)$ is defined as:

$$I_{2 \rightarrow 1}(\omega) = \log\left(\frac{S_{11}(\omega)}{\bar{H}_{11}(\omega)\Sigma_{11}\bar{H}_{11}^*(\omega)}\right), \quad (24)$$

and analogously, $I_{1 \rightarrow 2}(\omega)$,

$$I_{1 \rightarrow 2}(\omega) = \log\left(\frac{S_{22}(\omega)}{\hat{H}_{22}(\omega)\Sigma_{22}\hat{H}_{22}^*(\omega)}\right). \quad (25)$$

The instantaneous causality index $I_{1 \leftrightarrow 2}(\omega)$ is defined as:

$$I_{1 \leftrightarrow 2}(\omega) = \log\frac{(\bar{H}_{11}(\omega)\Sigma_{11}\bar{H}_{11}^*(\omega))(\hat{H}_{22}(\omega)\Sigma_{22}\hat{H}_{22}^*(\omega))}{\det(\mathbf{S}(\omega))}. \quad (26)$$

In equations (24) to (26), we have one index for each value ω of the frequency. Conversely, in the time domain

there was a single index for the GC between the two signals X_1 and X_2 . Just as discussed in Section 3, the indices $I_{2 \rightarrow 1}(\omega)$, $I_{1 \rightarrow 2}(\omega)$ and $I_{1 \leftrightarrow 2}(\omega)$ are greater than zero if there is any relation between the time series. They are zero otherwise.

Just like in the time domain, the total GC in the frequency domain is the sum of its individual components:

$$\begin{aligned} I(\omega) &= I_{2 \rightarrow 1}(\omega) + I_{1 \rightarrow 2}(\omega) + I_{1 \leftrightarrow 2}(\omega), \\ &= \log\left(\frac{S_{11}(\omega)S_{22}(\omega)}{\det(\mathbf{S}(\omega))}\right). \end{aligned} \quad (27)$$

The total GC is related to the so-called coherence $C_{12}(\omega)$ between signals (see Section C of the Appendix):

$$I(\omega) = -\log(1 - C_{12}(\omega)). \quad (28)$$

Moreover, we recover the GC in time domain through [15, 34]:

$$F_{i \rightarrow j} = \frac{1}{\omega_f - \omega_0} \int_{\omega_0}^{\omega_f} I_{i \rightarrow j}(\omega) d\omega. \quad (29)$$

5. Estimating Granger causality from data

In the last two sections we have mathematically defined the GC in both time and frequency domains. Here, we discuss how to calculate GC. In Section 5.1, we address a non-parametric estimation method that involves computing the Fourier and wavelet transforms of $X_1(t)$, and $X_2(t)$ [42–44]. In Section D of the Appendix, we address the parametric estimation of GC, which involves fitting the signals $X_1(t)$, and $X_2(t)$ to auto-regressive models (Section 2).

5.1. Calculating GC through Fourier and Wavelet Transforms

Here, we will give a numerical example for calculating and interpreting GC using a nonparametric estimation approach based on Fourier and wavelet transforms [34]. Our example consists of calculating the spectral matrix $\mathbf{S}(\omega)$ through the Fourier transform of the signals. For two stationary⁴ signals $X_1(t)$ and $X_2(t)$, the i, j element of the spectral matrix in equation (21) is

$$S_{ij}(\omega) = \frac{\langle \tilde{X}_i(\omega)\tilde{X}_j^*(\omega) \rangle}{T}, \quad i = 1, 2 \text{ and } j = 1, 2, \quad (30)$$

where T is the total duration of the signal, $\tilde{X}_i(\omega)$ is the discrete Fourier transform of $X_j(t)$ (calculated by a

⁴Stationarity, by definition, refers to time shift invariance of the underlying process statistics, which implies that all its statistical moments are constant over time [45]. There are several types of stationarity. Here, the required stationarity conditions for defining power spectral densities are constant means and that the covariance between any two variables apart by a given time lag is constant regardless of their absolute position in time.

fast fourier transform, FFT, algorithm) and $\tilde{X}_j^*(\omega)$ is its complex conjugate.

The variable ω contains the values of the frequency in the interval $[0, f_{\max}]$ corresponding to where the FFT was calculated. If Δt is the sampling time interval of the original signals, then the sampling frequency is $f_s = 1/\Delta t$ and $f_{\max} = f_s/2$, whereas the frequency interval contains $n_\omega = 1 + T/(2\Delta t)$ points. Then, for m signals ($m = 2$ in our example), we have a total of n_ω spectral matrices \mathbf{S} of dimensions $m \times m$. Recall that the diagonal elements of \mathbf{S} are called the autospectra, whereas the other elements are called the cross-spectra.

The transfer matrix, $\mathbf{H}(\omega)$ and the covariance matrix $\mathbf{\Sigma}$ are given by the decomposition of $\mathbf{S}(\omega)$ into the product of equation (20). The Wilson algorithm [34, 46, 47] (see also Section E of the Appendix) may be used for the decomposition of spectral matrices.

After determining these two matrices, we may calculate the GC indices through the direct application of equations (24) to (26).

For example, consider the autoregressive system studied in Ref. [34], which is given by:

$$\begin{aligned} X_1(t) &= 0.55X_1(t-1) - 0.8X_1(t-2) + CX_2(t-1) + \epsilon_1(t), \\ X_2(t) &= 0.55X_2(t-1) - 0.8X_2(t-2) + \epsilon_2(t). \end{aligned} \tag{31}$$

Here, $X_1(t)$ and $X_2(t)$ are $AR(2)$. The variable t is the time step index, such that the actual time is $t' = t \Delta t = t/f_s$. Besides, we know by construction that $X_2(t)$ influences $X_1(t)$ through the coupling constant C (although the opposite does not happen). The terms $\epsilon_1(t)$ and $\epsilon_2(t)$ are defined to have variance $\Sigma_{11} = \Sigma_{22} = 1$ and covariance $\Sigma_{12} = 0$ (they are independent random processes). To obtain a smooth power spectrum, we simulated 5000 trials of the system in equation (31) and computed the average spectra across trials. We set the parameters as $C = 0.25$, $f_s = 200$ Hz and $T = 25$ s, resulting in 5000 data points.

When $C = 0$, $X_1(t) = X_2(t)$, both processes are independent, and oscillate mainly in 40 Hz (Fig. 3a). For $C > 0$, the process X_1 receives input from X_2 , generating a *causal* coupling that is captured by the GC index in equations (24) and (25): a peak in 40 Hz in $I_{2 \rightarrow 1}$ indicates that process 2 (which oscillates in 40 Hz) is entering process 1 in this very frequency (Fig. 3c). The flat spectrum of $I_{1 \rightarrow 2}$ indicates that, on the other hand, process 2 does not receive input from 1. The absolute value of C changes the intensity of the GC peak. The instant causality index, $I_{2 \leftrightarrow 1}(\omega) = 0$ from equation (26), because $\Sigma_{12} = 0$ for all ω . The total GC in the system is obtained from the spectral coherence, equation (28). However, only the specific GC index reveal the directionality of the inputs between 1 and 2.

This simple example illustrates the meaning of causality in the GC framework: a Granger causal link is present if a process runs under the temporal influence of the past of another signal. We could have assumed C as a

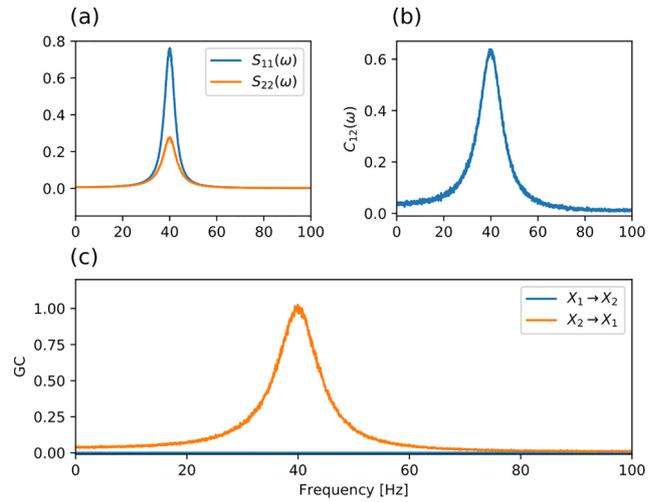


Figure 3: GC of a pair of autoregressive processes. GC for the system given in equation (31): by construction, the process 2 *causes* 1 by providing it input through the coupling constant $C = 0.25$. Parameters: total time $T = 25$ s and sampling frequency $f_s = 200$ Hz, resulting in 5000 time steps. **a.** Spectral matrix components calculated via equation (30). **b.** Coherence between signals 1 and 2, equation. (28). **c.** GC from 2 to 1 and 1 to 2, equation. (24) and (25): a peak in 40 Hz in the $I_{2 \rightarrow 1}$ GC index indicates that 2 Granger-causes 1, whereas the flat zero $I_{1 \rightarrow 2}$ shows, as expected, that 1 does not influence 2. The peak is in 40 Hz because process 2 has its main power in this frequency (see panel a).

time-varying function, $C(t)$, or even different parameters for the autoregressive part of each process alone; or the processes 1 and 2 could have been of different orders, implying in complex individual power spectra. These scenarios are more usual for any real world application [43, 48]. Then, instead of observing a clear peak for the GC indices, we could observe a more complex pattern with peaks that vary in time.

Instead of using the Fourier transform (which yields a single static spectrum for the whole signal), we may use the Wavelet transform [43, 49, 50] to yield time-varying spectra [51, 52]. Then, the auto and cross-spectra from equation (30) may be written as

$$S_{ij}(t, \omega) = \frac{\langle W_i(t, \omega) W_j^*(t, \omega) \rangle}{T}, \tag{32}$$

where $W_i(t, \omega)$ is the Wavelet transform of $X_i(t)$ and $W_i^*(t, \omega)$ is its complex conjugate. To compute the Wavelet transform, we use a Morlet kernel [43], with scale $s = 6$ oscillation cycles within a wavelet – a typical value for this parameter [53]. Similarly to what we did to the power spectrum, we measure the wavelet transforms for 5000 trials of the system in equation (31) in order to average the results. It is important to stress that a wavelet transform is applicable in this case because ensemble averages are being taken. Otherwise, estimates would be too unreliable for any meaningful inference.

In practice, there is one matrix \mathbf{S} for each pair (t, ω) ; or more intuitively, we have $n_T = T/\Delta t$ matrices $\mathbf{S}(\omega)$, each one for a given time step t . The decomposition of $\mathbf{S}(\omega)$ in equation (20) is done through Wilson’s algorithm. Then, we may calculate GC’s indices via equations (24) to (26) for each of the $\mathbf{S}(t, \omega)$ matrices with fixed t . This calculation results in $I_{2 \rightarrow 1}(\omega)$, $I_{1 \rightarrow 2}(\omega)$ and $I_{1 \leftrightarrow 2}(\omega)$ for each time step t . Finally, we concatenate these spectra across the temporal dimension, yielding $I_{2 \rightarrow 1}(t, \omega)$, $I_{1 \rightarrow 2}(t, \omega)$ and $I_{1 \leftrightarrow 2}(t, \omega)$.

For example, consider the same set of processes in equation (31), but with time-varying $C(t) = 0.25 H(t_0 - t)$, where $H(x) = 1$ for $x \geq 0$ (zero otherwise) is the Heaviside step function. The parameter t_0 is the time step index in which the coupling from 2 to 1 is turned off. This scenario is equivalent to having a set of concatenated constant C processes, such that the processes with $t > t_0$ have $C = 0$. Then, we expect the analysis in Fig. 3 to be valid for all the time steps $t < t_0$, and no coupling should be detected whatsoever for $t > t_0$.

That is exactly what is shown in Fig. 4: a sharp transition in the $I_{2 \rightarrow 1}(t, \omega)$ happens exactly at $t = t_0$ when C is turned off. The index $I_{1 \rightarrow 2}(t, \omega)$ remains zero for all the simulation. Again, this illustrates the meaning of GC in our system: whenever there is a directional coupling from a variable to another, there is nonzero GC in that link, in the example from signal $X_2(t)$ to $X_1(t)$.

6. Conditional Granger Causality

The concepts developed so far may be applied to a case with m variables. In this case, in order to try and infer

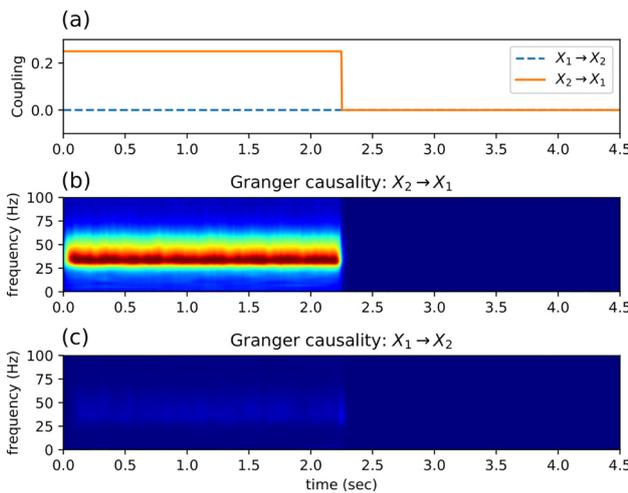


Figure 4: Time-varying GC in the frequency domain. GC of the system defined in equation (31), but with time-varying $C(t) = 0.25 H(t_0 - t)$. The spectral matrix is calculated via a Wavelet transform, equation (32), and decomposed for each time step t , yielding a temporal decomposition of the frequencies of the signals. **a.** Coupling constant as function of time. **b.** GC index from 2 to 1, $I_{2 \rightarrow 1}(t, \omega)$. **c.** GC index from 1 to 2, $I_{1 \rightarrow 2}(t, \omega)$.

the directionality⁵ of the interaction between two signals, in a system with m signals, we may use the so-called conditional Granger causality (cGC) [15, 33, 54, 55]. The idea is to infer the GC between signals i and j given the knowledge of all the other $m - 2$ signals of the system. This is done by comparing the variances obtained considering only i and j to the variances obtained considering all the other signals in the system. The AR model from Eqs (4) and (5) ends up having a total of m variables.

We may write the cGC in time domain as

$$F_{i \rightarrow j | k, \dots, m}, \tag{33}$$

or in the frequency domain as

$$I_{i \rightarrow j | k, \dots, m}(\omega). \tag{34}$$

But one may ask: “isn’t it simpler to just calculate the standard GC between every pair of signals in the system, always reducing the problem to a two-variable case?”

To answer that question, consider the case depicted in Fig. 5a: node 1 ($X_1(t)$) sends input to node 2 ($X_2(t)$) with a delay δ_{12} and sends input to node 3 ($X_3(t)$) with a delay δ_{13} . Measuring the pairwise GC between $X_2(t)$ and $X_3(t)$ suggests the existence of a coupling between them even if it does not physically exist (as in Fig. 5b). This occurs because signals $X_2(t)$ and $X_3(t)$ are correlated due to their common input from $X_1(t)$, and the simple pairwise GC between $X_2(t)$ and $X_3(t)$ fails to represent the correct relationship between the three nodes of Fig. 5a. The cGC solves this issue by considering the contribution of a third signal ($X_1(t)$ on this example) onto the analyzed pair ($X_2(t)$ and $X_3(t)$), as described below.

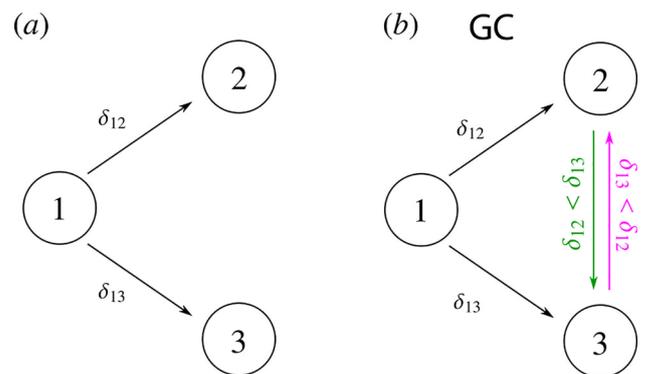


Figure 5: A system that GC fails to describe. **a.** Node 1 ($X_1(t)$) sends input to node 2 ($X_2(t)$) with delay δ_{12} and to node 3 ($X_3(t)$) with delay δ_{13} . **b.** A simple GC calculation wrongly infer a link from $X_2(t)$ to $X_3(t)$ if $\delta_{12} < \delta_{13}$, or from $X_3(t)$ to $X_2(t)$ if $\delta_{13} < \delta_{12}$. These links are not physically present in the system and appear only due to the cross-correlation between $X_2(t)$ and $X_3(t)$ caused by the common input $X_1(t)$.

⁵Whether i Granger-causes j or vice-versa.

To describe the system in (Fig. 5), equation (19) may be written as

$$\begin{pmatrix} X_1(\omega) \\ X_2(\omega) \\ X_3(\omega) \end{pmatrix} = \begin{pmatrix} H_{11}(\omega) & H_{12}(\omega) & H_{13}(\omega) \\ H_{21}(\omega) & H_{22}(\omega) & H_{23}(\omega) \\ H_{31}(\omega) & H_{32}(\omega) & H_{33}(\omega) \end{pmatrix} \begin{pmatrix} \epsilon_1^*(\omega) \\ \epsilon_2^*(\omega) \\ \epsilon_3^*(\omega) \end{pmatrix}, \quad (35)$$

where $X_3(t)$ has the noise term $\epsilon_3(t)$ with variance Σ_{33} . The corresponding spectral matrix $\mathbf{S}(\omega)$ is

$$\mathbf{S}(\omega) = \begin{bmatrix} S_{11}(\omega) & S_{12}(\omega) & S_{13}(\omega) \\ S_{21}(\omega) & S_{22}(\omega) & S_{23}(\omega) \\ S_{31}(\omega) & S_{32}(\omega) & S_{33}(\omega) \end{bmatrix}, \quad (36)$$

and the noise covariance matrix is

$$\mathbf{\Sigma} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{bmatrix}. \quad (37)$$

We want to calculate the cGC from $X_2(t)$ to $X_3(t)$ given $X_1(t)$, *i.e.* $F_{2 \rightarrow 3|1}$ in the time domain and $I_{2 \rightarrow 3|1}(\omega)$ in the frequency domain. The first step is to build a partial system from equation (35) ignoring the coefficients related to the probe signal $X_2(t)$, resulting in the partial spectral matrix $\mathbf{S}^p(\omega)$:

$$\mathbf{S}^p(\omega) = \begin{bmatrix} S_{11}(\omega) & S_{13}(\omega) \\ S_{31}(\omega) & S_{33}(\omega) \end{bmatrix}. \quad (38)$$

From this partial system, we can calculate $\mathbf{S}^p(\omega)$ and $\mathbf{S}(\omega)$ using the nonparametric methods already discussed above. Suppose that for $\mathbf{S}(\omega)$, we obtain the transfer matrix $\mathbf{H}(\omega)$ and the covariance matrix $\mathbf{\Sigma}$ (equation (37)), whereas for $\mathbf{S}^p(\omega)$ we obtain the transfer matrix $\mathbf{G}(\omega)$ and the covariance matrix $\mathbf{\rho}$:

$$\mathbf{\rho} = \begin{bmatrix} \rho_{11} & \rho_{13} \\ \rho_{31} & \rho_{33} \end{bmatrix}. \quad (39)$$

The matrices $\mathbf{H}(\omega)$ and $\mathbf{\Sigma}$ are 3×3 . The matrices $\mathbf{G}(\omega)$ and $\mathbf{\rho}$ are always one dimension less than the original ones, because they are built from the leftover rows and columns of the original system without the coefficients of the probe signal.

In the time domain, $F_{2 \rightarrow 3|1}$ is defined as

$$F_{2 \rightarrow 3|1} = \log \left(\frac{\rho_{33}}{\Sigma_{33}} \right), \quad (40)$$

or, in general,

$$F_{i \rightarrow j|k} = \log \left(\frac{\rho_{jj}}{\Sigma_{jj}} \right), \quad (41)$$

which is used to calculate the cGC from i to j given k , in time domain. Note that if the link between i and j is totally mediated by k , $\rho_{jj} = \Sigma_{jj}$, yielding $F_{i \rightarrow j|k} = 0$. However, the standard GC between i and j would result in a link between these variables. For our example in Fig. 5, we obtain $F_{2 \rightarrow 3|1} \gtrsim 0$, meaning that the influence

of $X_2(t)$ to $X_3(t)$ is conditioned on signal $X_1(t)$, and hence is almost null.

In the frequency domain, we first must define the transfer matrix $\mathbf{Q}(\omega) = \mathbf{G}(\omega)^{-1}\mathbf{H}(\omega)$. However, the dimensions of matrix $\mathbf{G}(\omega)$ do not match the dimensions of matrix $\mathbf{H}(\omega)$. To fix that, we add rows and columns from an identity matrix to the rows and columns that were removed from the total system in equation (35) when we built the partial system (*i.e.* we add the identity rows and columns to the rows and columns corresponding to signal $X_2(t)$ that was removed for generating $\mathbf{S}^p(\omega)$), such that:

$$\mathbf{G}(\omega) = \begin{bmatrix} G_{11}(\omega) & G_{13}(\omega) \\ G_{31}(\omega) & G_{33}(\omega) \end{bmatrix} \Rightarrow \begin{bmatrix} G_{11}(\omega) & 0 & G_{13}(\omega) \\ 0 & 1 & 0 \\ G_{31}(\omega) & 0 & G_{33}(\omega) \end{bmatrix}. \quad (42)$$

We can now safely calculate $\mathbf{Q}(\omega) = \mathbf{G}(\omega)^{-1}\mathbf{H}(\omega)$, from where we obtain $I_{2 \rightarrow 3|1}(\omega)$:

$$I_{2 \rightarrow 3|1}(\omega) = \log \left(\frac{\rho_{11}}{|Q_{11}(\omega)\Sigma_{11}Q_{11}^\dagger(\omega)|} \right), \quad (43)$$

or, in general,

$$I_{i \rightarrow j|k}(\omega) = \log \left(\frac{\rho_{jj}}{|Q_{jj}(\omega)\Sigma_{jj}Q_{jj}^\dagger(\omega)|} \right). \quad (44)$$

To illustrate the procedures for determining cGC, consider the system defined in Fig. 6 composed of 5 inter-

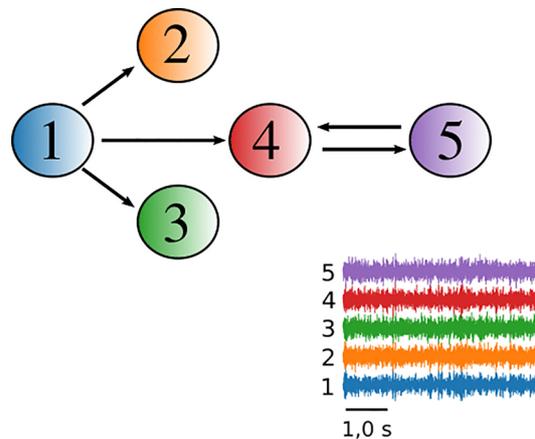


Figure 6: Many interacting components system. Illustration of a system with 5 interacting signals, having physical relations between them. We want to check whether GC or cGC is capable of capturing these interactions.

acting elements [26]⁶:

$$\begin{aligned}
 X_1(t) &= 0.95\sqrt{2}X_1(t-1) - 0.9025X_1(t-2) + \epsilon_1(t) \\
 X_2(t) &= 0.5X_1(t-2) + \epsilon_2(t) \\
 X_3(t) &= -0.4X_1(t-3) + \epsilon_3(t) \\
 X_4(t) &= -0.5X_1(t-2) + 0.25\sqrt{2}X_4(t-1) + \\
 &\quad + 0.25\sqrt{2}X_5(t-1) + \epsilon_4(t) \\
 X_5(t) &= -0.25\sqrt{2}X_4(t-1) + 0.25\sqrt{2}X_5(t-1) + \epsilon_5(t).
 \end{aligned}
 \tag{45}$$

Here, $X_1(t)$ sends its signal to $X_2(t)$, $X_3(t)$ and $X_4(t)$ with coupling intensities 0.5, -0.4 and -0.5 , respectively. Also, $X_4(t)$ sends input to $X_5(t)$ and *vice-versa* with couplings $-0.25\sqrt{2}$ and $0.25\sqrt{2}$ respectively. Note that X_1 sends signals to X_2 and X_4 with 2 time steps of delay, and to X_3 with 3 time steps of delay. X_4 and X_5 exchange signals with only 1 time step of delay.

Calculating the cGC index through equations (41) and (44), we recover the expected structure of the network (Figs. 7a and 8, respectively). The gray shades in Fig. 7 and the amplitude of the peaks in Fig. 8 are proportional to the coupling constants between each pair of elements. For comparison, Fig. 7b shows the simple pairwise GC, which detects connections that are not physically present in the system. Again, this occurs because the hierarchy of the network generates correlations between many pairs of signals that are not directly connected, as discussed in the example of Fig. 5.

It is important to note that the cGC connectivity not always reflects the underlying physical (or structural) connectivity between elements [35]. The example system in Fig. 6 is an illustrative simple case in which we obtained a neat result. However, real-world applications, such as inferring neuronal connectivity from brain signals, result in a cGC matrix that is more noisy due to multiple incoming signals and multiple delays. Thus, cGC is most generally referred to as giving “functional” connectivity, instead of structural connectivity.

7. Conclusion

Granger causality is becoming increasingly popular as a method to determine the dependence between signals in many areas of science. We presented its mathematical formulation and showed examples of its applications in general systems of interacting signals. This article also gives a contemporary scientific application of the Fourier transform – a subject that is studied in theoretical physics courses, but usually lacks practical applications in the classroom. We also used wavelet transforms, which may motivate students to learn more about the decomposition of signals in time and frequency domain, and its limitations through the uncertainty principle.

⁶Baccalá and Sameshima [26] analysed this system using partial directed coherence and directed transfer function.

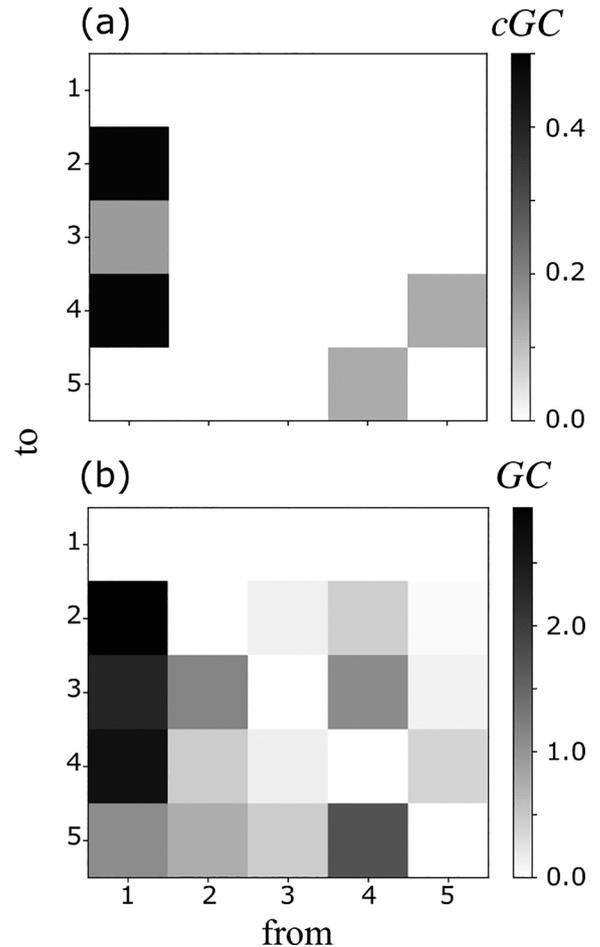


Figure 7: cGC and GC matrices in the time domain for a system of many interacting components. The system is given by equation (45) and is depicted in Fig. 6. **a.** Conditional GC, equation (41), captures exactly the physical interactions of the system. **b.** Simple pairwise GC, equation (29), captures the interactions, but also captures underlying correlations coming from the hierarchy of the network. Real systems often do not have a clear cGC matrix as given in panel **a** due to second order effects. The covariance matrices Σ and ρ were calculated using the nonparametric method.

We showed numerical examples, and explained them in an algorithmic way. We included the inference of steady and time-varying coupling, and the inference of connectivity in hierarchical networks via the conditional GC. A limitation of the GC is that it is ultimately based on linear regression models of stochastic processes (the AR models introduced in Section 2). Other measures, such as the transfer entropy, are more suitable to describe nonlinear interactions, and do not need to be fitted to an underlying model. Even though the nonparametric estimation of GC does not rely on fitting, it is still a measure of linear interaction. It is also possible to show that both GC and transfer entropy yield the same results for Gaussian variables [56].

In spite of the existing debate about what exactly the GC captures, specially in the neuroscience commu-

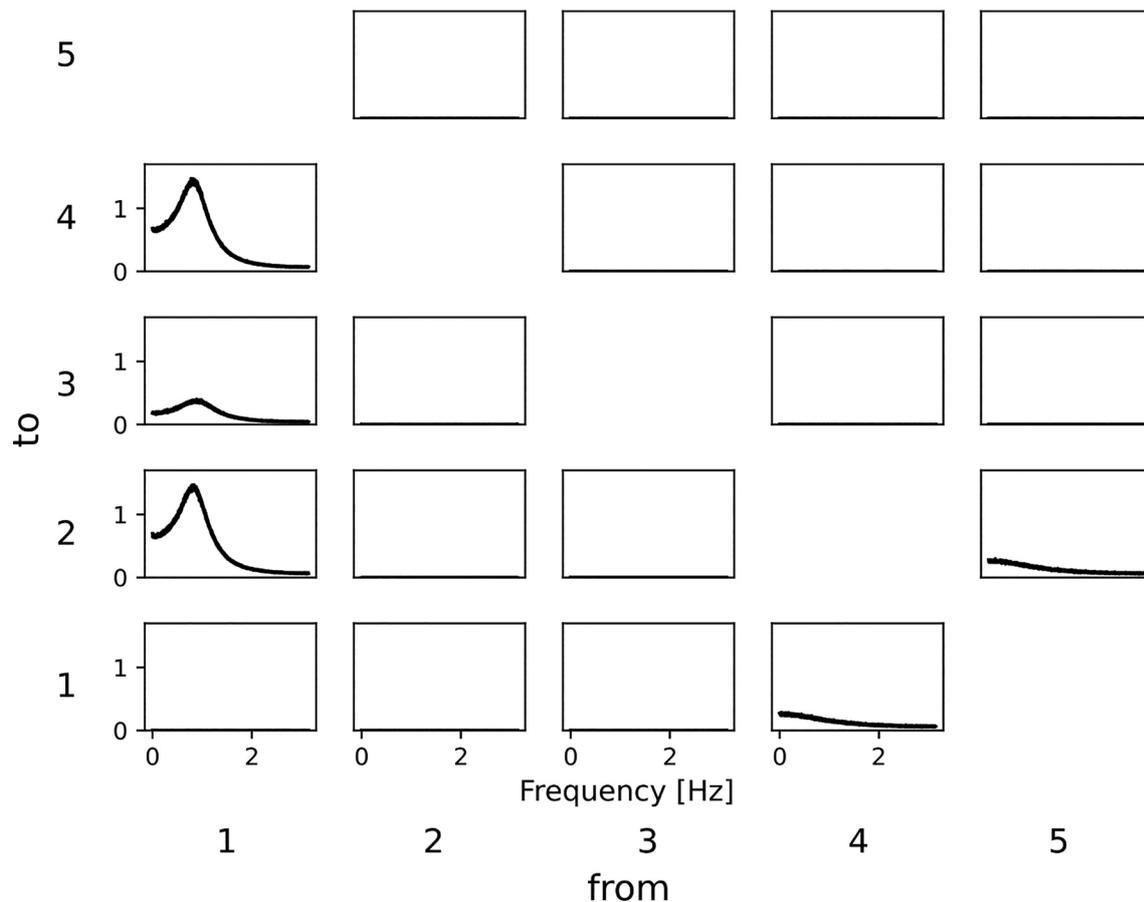


Figure 8: cGC and GC matrices in the frequency domain for a system of many interacting components. The same system as depicted Fig. 6. Each panel in row i and column j corresponds to the cGC index between elements i and j in the frequency domain, equation (44).

nity [57, 58], GC has become a well-established measurement for the flux of information in the nervous system [59]. And here, we hope to have provided the necessary tools to those who wish to learn the basic principles and applications underlying GC.

Acknowledgments

This article was produced as part of the S. Paulo Research Foundation (FAPESP) Research, Innovation and Dissemination Center for Neuromathematics (CEPID NeuroMat, Grant No. 2013/07699-0). The authors also thank FAPESP support through Grants No. 2013/25667-8 (R.F.O.P.), 2015/50122-0 (A.C.R.), 2016/03855-5 (N.L.K.), 2017/07688-9 (R.O.S), 2018/20277-0 (A.C.R.) and 2018/09150-9 (M.G.-S.). V.L. is supported by a CAPES PhD scholarship. A.C.R. thanks financial support from the National Council of Scientific and Technological Development (CNPq), Grant No. 306251/2014-0. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Supplementary material

The following online material is available for this article:

- Apendix A
- Apendix B
- Apendix C
- Apendix D
- Apendix E

All codes were developed in Python and are available in: <https://github.com/ViniciusLima94/pyGC>.

References

- [1] A. Falcon, in: *The Stanford Encyclopedia of Philosophy*, edited by E.N. Zalta (Stanford University, Palo Alto, 2019).
- [2] C.W.J Granger, *Econometrica* **37**, 424 (1969).
- [3] N. Wiener, in: *Modern Mathematics for Engineers* edited by E.F. Beckenbach (McGraw-Hill, New York, 1956), v. 1.
- [4] J. Pearl, *Biometrika* **82**, 669 (1995).
- [5] J.Y Halpern and J. Pearl, *Br. J. Philos. Sci.* **56**, 843 (2005).
- [6] J. Halpern, in: *24th International Joint Conference on Artificial Intelligence* (Buenos Aires, 2015).

- [7] T. Vÿrost, Š. Lyócsa and E. Baumöhl, *Physica A*. **427**, 262 (2015).
- [8] B. Candelson and S. Tokpavi, *J. Bus. Econ. Stat.* **34**, 240 (2016).
- [9] H. Ding, H. Kim and S.Y. Park, *Energ. Econ.* **59**, 58 (2016).
- [10] J. Runge, S. Bathiany, E. Bollt, G. Camps-Valls, D. Coumou, E. Deyle, C. Glymour, M. Kretschmer, M.D. Mahecha, J. Muñoz-Marí et al., *Nat. Commun.* **10**, 1 (2019).
- [11] D.A. Smirnov and I.I. Mokhov, *Phys. Rev. E*. **80**, 016208, (2009).
- [12] P.O. Amblard and O.J.J. Michel, *Entropy* **15**, 113 (2013).
- [13] G. Tissot, A. Lozano-Durán, L. Cordier, J. Jiménez and B.R. Noack, *J. Phys. Conf. Ser.* **506**, 012006 (2014).
- [14] A. Brovelli, M. Ding, A. Ledberg, Y. Chen, R. Nakamura and S.L. Bressler, *P. Natl. A. Sci.* **101**, 9849 (2004).
- [15] M. Ding, Y. Chen and S.L. Bressler, in: *Handbook of Time Series Analysis: Recent Theoretical Developments and Applications*, edited by B. Schelter, M. Winterhalder and J. Timmer (Wiley-VCH, Weinheim, 2006).
- [16] F.S. Matias, L.L. Gollo, P.V. Carelli, S.L Bressler, M. Copelli and C.R. Mirasso, *Neuroimage* **99**, 411 (2014).
- [17] S. Hu, H. Wang, J. Zhang, W. Kong, Y. Cao and R. Kozma, *IEEE T. Neur. Net. Lear.* **27**, 1429 (2015).
- [18] A.K. Seth, A.B. Barrett and L. Barnett, *J. Neurosci.* **35**, 3293 (2015).
- [19] M. Havlicek, J. Jan, M. Brazdil and V.D. Calhoun, *Neuroimage* **53**, 65 (2010).
- [20] W. Liao, D. Mantini, Z. Zhang, Z. Pan, J. Ding, Q. Gong, Y. Yang and H. Chen, *Biol. Cybern.* **102**, 57 (2010).
- [21] L. Pollonini, U. Patidar, N. Situ, R. Rezaie, A.C. Papanicolaou and G. Zouridakis, in: *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology* (Buenos Aires, 2010).
- [22] A.B. Barrett, M. Murphy, M.A. Bruno, Q. Noirhomme, M. Boly, S. Laureys and A.K. Seth, *PLoS ONE* **7**, e29072 (2012).
- [23] M.E. Lynall, D.S. Bassett, R. Kerwin, P.J. McKenna, M. Kitzbichler, U. Muller and E. Bullmore, *J. Neurosci.* **30**, 9477 (2010).
- [24] S.H. Jin, P. Lin, S. Auh and M. Hallett, *Movement Disord.* **26**, 1274 (2011).
- [25] V. Lima, R.F.O. Pena, C.A.C. Ceballos, R.O. Shimoura and A.C. Roque, *Rev. Bras. Ensino. Fis.* **41**, e20180197 (2019).
- [26] L.A. Baccalá and K. Sameshima, *Biol. Cybern.* **84**, 463 (2001).
- [27] P. Fries, *Trends Cogn. Sci.* **9**, 474 (2005).
- [28] M.J. Kamiński and K.J. Blinowska, *Biol. Cybern.* **65**, 203 (1991).
- [29] D. La Rocca, P. Campisi, B. Vegso, P. Cserti, G. Kozmann, F. Babiloni and F. De Vico Fallani, *IEEE T. Bio-Med. Eng.* **61**, 2406 (2014).
- [30] T. Schreiber, *Phys. Rev. Lett.* **85**, 461 (2000).
- [31] R. Vicente, M. Wibral, M. Lindner and G. Pipa, *J. Comput. Neurosci.* **30**, 45 (2011).
- [32] J. Geweke, *J. Am. Stat. Assoc.* **77**, 304 (1982).
- [33] J.F. Geweke, *J. Am. Stat. Assoc.* **79**, 907 (1984).
- [34] M. Dhamala, G. Rangarajan and M. Ding, *Phys. Rev. Lett.* **100**, 018701 (2008).
- [35] A. Seth, *Scholarpedia*. **2**, 1667 (2007).
- [36] D. Marinazzo, W. Liao, H. Chen and S. Stramaglia, *Neuroimage* **58**, 330 (2011).
- [37] G.E.P. Box, G.M. Jenkins, G.C. Reinsel and G.M. Ljung, *Time Series Analysis: Forecasting and Control* (Wiley, Hoboken, 2015), 5^a ed.
- [38] G. Eshel, Internet resource **2**, 68 (2003).
- [39] W. Van Drongelen, *Signal Processing for Neuroscientists*, (Academic Press, London, 2018), 2^a ed.
- [40] R. Takalo, H. Hytti and H. Ihalainen, *J. Clin. Monit. Comput.* **19**, 401 (2005).
- [41] R. Meddins, *Introduction to Digital Signal Processing* (Newnes, Oxford, 2000), 1^a ed.
- [42] I. Daubechies, *Ten Lectures on Wavelets* (Society for Industrial and Applied Mathematics, Philadelphia, 1992).
- [43] C. Torrence and G.P. Compo, *Bull. Amer. Meteor. Soc.* **79**, 61 (1998).
- [44] P. Stoica and R.L. Moses, *Spectral Analysis of Signals* (Pearson Prentice Hall, Upper Saddle River, 2005).
- [45] W.A. Woyczyński, *A First Course in Statistics for Signal Analysis* (Birkhäuser, Boston, 2019), 3^a ed.
- [46] G.T. Wilson, *SIAM J. Appl. Math.* **23**, 420 (1972).
- [47] G.T. Wilson, *J. Multivariate Anal.* **8**, 222 (1978).
- [48] P. Holme and J. Saramäki, *Phys. Rep.* **519**, 97 (2012).
- [49] A.D. Poularikas, *Transforms and Applications Handbook* (CRC press, Florida, 2010).
- [50] P.S. Addison, *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance* (CRC press, Florida, 2017).
- [51] M.J.A. Bolzan, *Rev. Bras. Ensino. Fis.* **28**, 563 (2006).
- [52] M.O. Domingues, O. Mendes, M.K. Kaibara, V.E. Mencioni and E. Bernardes, *Rev. Bras. Ensino. Fis.* **58**, 228 (2016).
- [53] M. Farge, *Annual review of fluid mechanics* **24**, 395 (1992).
- [54] Y. Chen, S.L. Bressler and M. Ding, *J. Neurosci. Meth.* **150**, 228 (2006).
- [55] S. Malekpour and W.A. Sethares, *Biol. Cybern.* **109**, 627 (2015).
- [56] L. Barnett, A.B. Barrett and A.K. Seth, *Phys. Rev. Lett.* **103**, 238701 (2009).
- [57] P.A. Stokes and P.L. Purdon, *P. Natl. A. Sci.* **114**, E7063 (2017).
- [58] L. Barnett, A.B. Barrett and A.K. Seth, *Neuroimage* **178**, 744 (2018).
- [59] S.L. Bressler and A.K. Seth, *Neuroimage* **58**, 323 (2011).