PeerJ Preprints NOT PEER-REVIEWED

causal effects could be identified by G-causality analysis. These results demonstrate a good sensitivity and specificity of the conditional G-causality analysis in the time domain when applied on covariance stationary, non-correlated electrophysiological signals.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26703v1 | CC BY 4.0 Open Access | rec: 15 Mar 2018, publ: 15 Mar 2018

PeerJ Preprints NOT PEER-REVIEWED

The reliability of conditional Granger causality analysis in the time domain

Raffaella Franciotti[1,*], Nicola Walter Falasca[1,2]

[1]Department of Neuroscience, Imaging and Clinical Science, G. d'Annunzio University, Chieti, Italy
[2]BIND – Behavioral Imaging and Neural Dynamics Center, University of Chieti-Pescara, Chieti, Italy

*Corresponding author:
Raffaella Franciotti
Department of Neuroscience, Imaging and Clinical Sciences "G. d'Annunzio" University
Via Luigi Polacchi
66013 Chieti, Italy
Ph:+39 08713556901; Fax:+39 08713556930
e-mail: raffaella.franciotti@itab.unich.it

Raffaella Franciotti ORCID ID: 0000-0002-8819-1061

PeerJ Preprints NOT PEER-REVIEWED

Abstract

Background. Brain function requires a coordinated flow of information among functionally specialized areas. Quantitative methods provide a multitude of metrics to quantify the oscillatory interactions measured by invasive or non-invasive recording techniques. Granger causality (G-causality) has emerged as a useful tool to investigate the directions of information flows, but challenges remain on the ability of G-causality when applying on biological data. In addition it is not clear if G-causality can distinguish between direct and indirect influences and if G-causality reliability was related to the strength of the neural interactions.

Methods. In this study time domain G-causality connectivity analysis was performed on simulated electrophysiological signals. A network of 19 nodes was constructed with a designed structure of direct and indirect information flows among nodes, which we referred to as a ground truth structure. G-causality reliability was evaluated on two sets of simulated data while varying one of the following variables: the number of time points in the temporal window, the lags between causally interacting nodes, the connection strength between the links, and the noise.

Results. Results showed that the number of time points in the temporal window affects G-causality reliability substantially. A large number of time points could decrease the reliability of the G-causality results, increasing the number of false positive (type I errors). In the presence of stationary signals, G-causality results are reliable showing all true positive links (absence of type II errors), when the underlying structure has the delays between the interacting nodes lower than 100 ms, the connection strength higher to 0.1 time the amplitude of the driver signal and good signal to noise ratio. Finally, indirect links were revealed by G-causality analysis for connection strength higher than the direct link and lags lower than the direct link.

Discussion. Conditional multivariate vector autoregressive model was applied to 19 virtual time series to estimate the reliability of the G-causality analysis on the identification of the true positive link, on the presence of spurious links and on the effects of indirect links. Simulated data revealed that weak direct but not weak indirect causal effects could be identified by G-causality analysis. These results demonstrate a good sensitivity and specificity of the conditional G-causality analysis in the time domain when applied on covariance stationary, non-correlated electrophysiological signals.

Keywords: Granger causality; simulation; temporal window; lags; gain; noise.

Introduction

The characterization of the functional connectivity (FC) patterns of the human brain is a key challenge in neuroscience. FC is defined as the statistical dependence among measured time series usually evaluated in terms of correlations or mutual information. Recent advances have shown that FC patterns of human brain networks can be non-invasively characterized from a variety of electrophysiological (electroencephalography (EEG), magnetoencephalography) and neuroimaging techniques, e.g., structural, diffusion and functional MRI (fMRI). Specifically, resting state EEG is a promising technique to derive FC patterns in terms of non-invasive, cost-effective nature and independence of task difficulty. The hypothesis is that neuronal oscillations provide a mechanism underlying dynamic coordination in the brain (Varela et al., 2001; Fries, 2015). A multitude of metrics quantifies oscillatory interactions (see Wang et al., 2014 for a review). A first subdivision among methods can be made on whether the metric quantifies the direction of the interaction. Nondirected FC metrics seek to capture some form of interdependence between signals, without reference to the direction of influence (Bastos & Schoffelen, 2016). FC can be quantified by measures of statistical dependencies, such as correlations, coherence in the frequency domain, or transfer entropy (TE). It is not based on model comparison because the correlations are attributes of the data, not the model (Friston, 2011). In contrast, effective connectivity is based on the comparison of a model with and without a particular connection to infer its presence (Friston, 2011). Effective connectivity measures can distinguish the driver from the recipient estimating the direction of information flow and provide effective mechanisms estimating the interaction coefficients. Among the statistically principled techniques which estimate the direction of influence in time series, Granger causality (G-causality) has emerged in recent years as a useful tool to investigate the directions of neuronal interactions (Brovelli et al., 2004; Hesse et al., 2003) and is probably the most prominent and powerful technique (Nolte et al., 2008; Seth, Barrett & Barnett, 2015).

For fMRI acquisitions, G-causality has been applied to core regions recruited during a task of motor execution and motor imagery (Gao, Duan & Chen, 2011), during spatial working memory task in early deaf subjects (Ding et al., 2015), or during resting state (Franciotti et al., 2013). In epilepsy, G-causality analysis has been shown to depict pathway important for seizure propagation (Coben et al., 2015) or for diagnosis and monitoring purposes (Protopapa et al., 2016), G-causality analysis reported similar results to dynamic causal modelling applied to fMRI acquisition (David et al., 2008) and to pathways revealed by diffusion tensor imaging (Bhardwaj et al., 2010). Successful results of G-causality were also found on EEG data. Indeed it showed higher inter-subject consistency than the synchrony analysis during loss of consciousness (Barret et al., 2012). It classified ‘awake’ and ‘anesthetized’ states of patients with an high accuracy (Nicolau et al., 2012) and was able to distinguish patients with mild cognitive impairment from age-matched control subjects (Dauwels et al., 2010). In addition G-causality method was found to be useful to study also complex cognitive functions distinguishing distinct network patterns during visuo-spatial working memory task (Protopapa et al., 2014).

G-causality seeks to establish a statistical causation from the data that is based on the maxim that causes precede their effects. The statistical causality approach was originally developed by Wiener (1956) and later implemented using auto-regressive models by Granger (1969). These measures have recently been supplemented with methods like multivariate G-causality to provide measures that are less sensitive to indirect links (Barnett & Seth, 2013). The

G-causality calculates the directed connectivity (j→i) based on the notion that information in the past of j helps to predict the future of i with greater accuracy than by only considering the past of i itself (Granger, 1969). However, challenges remain on the ability of G-causality to infer the direction of information flows among nodes of a network when applying on biological data. In addition it is not clear if G-causality can distinguish between direct and indirect influences and if G-causality reliability was related to the strength of the neural interactions.

The aim of this study was to establish the reliability of G-causality when varying one of the different conditions: the length of the temporal window, the lag between causally interacting nodes, the strength of the connection, and the adding noise. By means of 19 simulated signals mimicking 19 channels of the international 10-20 EEG system, we assessed the reliability and robustness of G-causality results by means of parameters of sensibility and specificity related to the number of true positive and false positive links. In addition G-causality reliability was compared for direct and indirect link and was estimated how it was modulated by the strength of the connection.

Material and Methods

Simulated data

G-causality analysis can be performed under these assumptions:

1) data sets are not correlated, thus the data-generating processes in any time series have to be independent variables,

2) the signals are stationary, more specifically the signals must be covariance stationary (also known as weak or wide-sense stationarity, i.e., the means and variances of the time series are stable over time),

3) the model does not have any unit roots which are one cause for non-stationarity.

If the stationarity of the signals is not satisfied, used models for G-causality are invalid and may contain so-called 'spurious regression' results, i.e., correlations that arise from non-stationarities rather than from relations among signals (Granger and Newbold 1974).

Simulated signals *X(t)* were generated in order to guarantee these assumptions, summing the sinusoidal signals, with a random phase for each frequency *i* and according to the following formula:

$$X(t) = \sum_{i=1}^{n} A_i \sin(2\pi f_i t + \varphi_i)$$

where n=2460 (considering a sampling rate of 256 Hz and the frequency resolution of 0.05 Hz),

A_i and f_i the amplitude and frequency

$$\varphi_i = randn \cdot 2\pi$$

randn is a random scalar drawn from the standard normal distribution.

G-causality is not based on the phase differences between signals, thus the results were not influenced by the distortion of the signals by the randomization of the phases.

In addition, randomly generated white Gaussian noise was performed by means of the awgn(snr) function from matlab library and was super-imposed to each signal, obtaining a simulated raw signal for each channel. The scalar

parameter snr specifies the signal-to-noise ratio per sample, in dB. The simulated raw data are available as supplemental file (see S.mat).

The link from the signal *Y(t)* to the signal *X(t)* was simulated according to the following formula:

$$X(t+lag) = X(t+lag) + Amp \cdot Y(t)$$

where

$$t = 1.....(length(X) - lag)$$

The *lag* and *Amp* parameters were the delay and the connection strength factor in the information flow from the signal Y (driver) to the signal X (recipient).

In the simulation the designed structure of information flow did not change, while the length of the temporal window, the time delays (the lags) of the interactions across nodes, the connection strength and the noise were selected as variables. They changed one at a time in order to evaluate the effects of these variables on the reliability of G-causality. The range of the variations of these variables was chosen in order to obtain the model consistency of G-causality higher than 80%, as values below 80% may give cause for concern (Seth, 2010).

Two examples of simulated data were performed. The first example consisted of 19 nodes, with five sources (drivers), seven sinks (recipients) and seven links. The delay values (i.e. the lags between causal interacting nodes) and the gain values (i.e. the connection strength factor which multiplied the amplitude of the driver signal) between nodes are reported for the starting designed structure in Table 1. In the simulation, we added a value ranging from -9 to 30 time points to the lag values and a value ranging from -0.2 to 0.2 to the connection strength factor of the starting designed structure. In addition, to simulate the presence of an indirect link we added in the designed structure a latent 20th signal to simulate a connection from node 10 to node 20 and a connection from node 20 to node 13. The causal link from the node 10 to the node 20, which it is not included in the analysis, had the lag of 11 time points and the connection strength factor fixed to 0.5. The link from the node 20 to the node 13 had the starting value of the lag of 14 time points and the strength factor was performed according to the formula: 0.6+2·connecting strength factor (ranging from -0.2 to 0.2).

The aim was to verify if G-causality would pick the false link from 10 to 13 when the 20th signal was not included in the analysis.

To investigate the ability of the G-causality to distinguish weak direct causal effects, a second example of simulation was performed. To the designed structure of the first simulation, 7 links were added. Then, the new system of the second example consisted of 19 nodes with 12 sources, 11 sinks and 14 links. Table 2 shows the starting values for the designed structure of this second example. Moreover, the indirect link from node 10 to node 13 was also included with the same characteristics of the first example.

Figure 1 shows a schematic representation of the designed structures for the two examples. The Matlab code to simulate the designed structure while varying the number of time points in the temporal window, the lags between causally interacting nodes, the connection strength between the links, and the noise is available as Supplemental file for the first and the second example (see code.m).

G-causality analysis on simulated data

Time domain G-causality connectivity analysis was applied to identify patterns of causal interaction between nodes. According to linear vector autoregressive (VAR) models, two wide-sense stationary time series *X (t)* and *Y(t)* can be explained by their own past by means of a linear model with coefficients a_j and b_j and prediction errors $\varepsilon_1(t)$ and $\eta_1(t)$ respectively:

$$X(t)=\sum_{j=1}^{m} a_j X(t-j)+\varepsilon_1(t)$$

$$Y(t)=\sum_{j=1}^{m} b_j Y(t-j)+\eta_1(t)$$

Lagged vector autoregression models are used to determine the ability of one time-varying signal to predict the future behaviour of another, comparing the accuracy of the prediction obtained by considering only information of the own past than the inclusion of the past of the other signal of the system (Granger, 1969). Thus, the temporal dynamics of the time series *X(t)* and *Y(t)* (both of length T) can be described also including in the model not only information on the own past of the time series, but also information on the past of the other time series, with prediction errors $\varepsilon_2(t)$ and $\eta_2(t)$ which are different from the previous $\varepsilon_1(t)$ and $\eta_1(t)$.

$$X(t)=\sum_{j=1}^{m} a_j X(t-j)+\sum_{j=1}^{m} b_j Y(t-j)+\varepsilon_2(t)$$

$$Y(t)=\sum_{j=1}^{m} c_j Y(t-j)+\sum_{j=1}^{m} d_j X(t-j)+\eta_2(t)$$

where *m* is the maximum number of lagged observations included in the model (the model order, $m << T$), whereas b_j and d_j are the gain factors, respectively, of the signal *Y(t)* (driver) influencing the signal *X(t)* (recipient), and of the signal *X(t)* (driver) influencing the signal *Y(t)* (recipient).

The linear influence from *X(t)* to *Y(t)* ($F_{X\to Y}$) and from *Y(t)* to *X(t)* ($F_{Y\to X}$) can be calculated as the log ratio between the variances of the residual errors.

$$F_{X\to Y}=\log\left(\frac{\mathrm{var}(\eta_1)}{\mathrm{var}(\eta_2)}\right)$$

$$F_{Y\to X}=\log\left(\frac{\mathrm{var}(\varepsilon_1)}{\mathrm{var}(\varepsilon_2)}\right)$$

G-magnitude is given by the log ratio of the variance of the prediction-error terms for the reduced (omitting the signal of the potential cause) and full regressions (including the signal of the potential cause).

G-causality analysis is generalized to the multivariate (conditional) case in which the G-causality of *Y(t)* on *X(t)* is tested in the context of multiple additional variables (Geweke, 1982) when all other variables are also included in the regression model.

In our simulation study, conditional multivariate VAR (MVAR) model was applied to 19 time series (Seth, 2010). Data analysis was performed using the in-house software BSMART, a MATLAB/C Toolbox implemented to analyse brain circuits (Cui et al., 2008). MVAR model was applied to the 19 time series to estimate G-causality

connectivity (Seth, 2010). The method of ordinary-least-squares was used to compute the regression coefficients. The F-statistic, Bonferroni-corrected (nominal p value of 0.05, then divided for multiple comparisons by n·(n-1) where n=19), was applied to the coefficients of the MVAR model. In the presence of high dimensional time series, conditional G-causality could fail because of the large number of coefficients to be estimated. The problem of reducing the number of model coefficients has been addressed in the last years, developing many methods of variable subset selection. These methods include the simple sequential search method and stepwise methods implementing the bottom-up (forward selection) and top-down (backward elimination) strategies or more complicated schemes such as the genetic algorithms, the particle swarm optimization and the ant colony optimizations (Siggiridou & Kugiumtzis, 2016). Another method known as backward-in-time selection takes also into account the lag dependence structure, implementing a supervised stepwise forward selection guided by the lag order of the lagged variables (Vlachos & Kugiumtzis, 2013).

The Akaike information criterion (Akaike, 1974) was used to estimate the order of the model (Bressler & Seth, 2011). Covariance stationarity was checked by using the Durbin–Watson test, based on MATLAB code provided by Seth (Seth, 2010) and the Dickey–Fuller test ($p<0.01$) to identify unit roots. The consistency of the MVAR model, which ensures that the MVAR model properly represents the data, was verified by the tests proposed by Ding (Ding et al., 2000), whereas the Durbin–Watson statistics assessed whether the residuals are uncorrelated.

G-causality matrix of 19 rows and 19 columns represents the causal strength of the connection between each couple of nodes. The G-causality analysis was applied on a single time window because when multiple time windows are used the mean connectivity matrix which includes the mean connection strength across all time windows must be thresholded to produce a network graph. The choice of the most appropriate threshold value is an unresolved problem (Wang et al., 2014).

To estimate the reliability of G-causality we calculated the sensitivity and specificity of the results

Results

Length of the temporal window

The length of the temporal window of the time series used to perform G-causality is a crucial factor. It needs to be as short as possible since, in real datasets, FC may change dynamically over time, but it needs to be as long as possible to have reliable G-causality results. Indeed the length of the temporal window should be at least n·p, where n is the number of time series and p the model order (Seth, 2010). For the simulated data, we performed G-causality from 512 (2 s) to 10496 (41 s) time points. The time delays of the interactions between channels (lags), the connection strength (Table 1 for the first and Table 2 for the second example) and the noise did not change while varying the number of time points of the time series. For the window length from 512 time points to 1280 time points the model order of G-causality was higher than 40 time points for both the simulations and the results were not reliable. In the first example, for window length equal or greater than 1536 time points, all the causal information flows were revealed by the G-causality analysis (the sensitivity was always 1), but the number of false positive links was high (see Figure 2 for 1792 time points). The specificity reached the maximum value of 0.99 for the temporal window of

4096 time points. The false link from 10 to 13 was revealed for temporal window length higher than 4608 time points (see Figure 2 for 1792 and 4352 time points). The values of the model order increased over the window length with a range from 17 to 29 time points. The specificity ranged from 0.96 to 0.99. The model consistency was always higher than 90% over the window length, ranging from 92% to 95%.

The second example confirmed the results of the first simulation. The temporal window of 5120 time points was used for both the examples while varying lags between channels, connection strength factor and the noise separately because for this temporal window the false link from 10 to 13 was revealed by the G-causality analysis.

Lags and model order

The signalling between neurons and brain regions involves time delays, which must be taken into consideration in G-causality analysis. To assess the effect of lags between nodes on G-causality reliability, we varied the lags of the designed matrix adding a value of all lag values reported in Table 1 for the first example and in Table 2 for the second, example. This adding value ranged from -9 to 30 time points with a step of 1. Specifically, when we subtracted 9 to the starting designed structure in which the lags ranged from 10 to 16 time points across channels pairs (see Table 1 and Table 2), we obtained a structure with lags ranged from 1 to 7 time points. Other parameter values were 5120 time points (20s) for the length of the time series and the factor of the connection strength of the causal link as reported in Table 1 and Table 2.

For the first example when the lags of the designed structure ranged from 1 to 7 time points (the adding value for the lags was -9), the sensitivity was 1 indicating that all the causal information flows were detected by the G-causality analysis, and the number of false positive was low (the specificity was 0.99, see Figure 3 for -9). The sensitivity was always 1 until the adding value for the lags was 24 (the lags of the designed structure ranged from 34 to 40 time points). Thus, for the designed structure, G-causality reached the best concordance with the designed structure (see Figure 3 for adding value of 5) when the lags across nodes ranged from 15 time points (59 ms) to 21 time points (82 ms). For adding values of lags higher than 24 (the lags across nodes were higher than 34 time points), the sensitivity decreased drastically (see Figure 3 for 29). The values of the model order increased over the lags from 23 to 40 time points, whereas the model consistency was always around 94%.

The indirect link from 10 to 13 was revealed when the adding values of the lags of the designed structure ranged from -9 (lags between signals from 1 to 7 time points) to 0 (lags between signals from 10 to 16 time points).

For the second example, similar results were found. Figure 4 shows the graphs of the results for lags ranging from 3 to 9 time points (adding value to the starting structure was -7), ranging from 10 to 16 time points (no adding value), ranging from 39 to 45 time points (adding value of 29). When the lags across nodes were higher than 39 time points the sensitivity decreased drastically (Figure 4 for 29). When the designed structure had lag values between nodes ranging from 1 to 19 time points (the adding value for the lags ranged from -9 to 3) the model order was 20. The maximum model order was 35 for lags of the designed structure ranged from 32 to 38 time points (the adding value for the lags was 22). The indirect link from 10 to 13 was revealed when the adding values of the lags of the designed structure ranged from -9 (lags between signals from 1 to 7 time points) to -1 (lags between signals from 9 to 15 time points).

Influence of connection strength

To evaluate the effect of the connection strength on G-causality reliability, we varied the connection strength factor which multiplied the amplitude of the driver signal shown in Table 1 and in Table 2. A value which ranged from -0.2 to 0.2 with a step of 0.01 was added to each connection strength factor of all the causal links for the first example (Table 1) and to connection strength factor with an "*" in Table 2 for the second example. Specifically, to the starting designed structure in which the connection strength factor ranged from 0.3 to 0.55 (first example, Table 1) and from 0.05 to 0.55 (second example, Table 2), we subtracted 0.2 and we obtained a structure with connection strength factor ranged from 0.1 to 0.35 (first example) and from 0 to 0.35 (second example). Other parameter values were 5120 time points (20s) for the length of the time series and the lags as reported in the Table 1 and 2. All these parameters as well as the noise did not change while varying the connection strength.

In the first example, when the connection strength between nodes ranged from 0.1 to 0.35 the sensitivity was 0.86, indicating that G-causality analysis was not able to identify all the causal information flows, whereas the number of false positive was low (the specificity was 0.99, see Figure 5 for -0.2). The sensitivity was 1 when the connection strength factor of the structure was higher than 0.17 (adding value of -0.13). From adding factor ranging from 0.01 to 0.07 the specificity was highest, then it decreased (see Figure 5 for 0.05 and 0.1). The values of the model order ranged from 22 to 25 time points and the model consistency increased linearly over connection strength even if only little (from 93.8% to 94.1%). The indirect link from 10 to 13 was revealed when the connection strength factor was higher than 0.33 (the adding value to the connection strength factor was 0.03). The influence of the connection strength was higher on the direct than the indirect link.

In the second example the link from node 3 to node 6 with connection strength factor fixed to 0.05 was not revealed, whereas the link from node 7 to node 19 with connection strength factor fixed to 0.1 was revealed until the other links of the designed structure had connection strength ranging from 0.05 to 0.55. The link from node 19 to node 3 with connection strength factor fixed to 0.15 was always revealed. The link from node 1 to node 3 was revealed when its connection strength was higher or equal to 0.1 (adding value of -0.1). The indirect link from node 10 to node 13 was instead revealed when the connection strength from node 20 to node 13 was higher or equal to 0.5 (adding value of 0.05). Figure 6 shows the graphs of the results for adding value of the connection strength factor from -0.2 to 0.15 with a step of 0.05. The values of the model order ranged from 20 to 25 time points and the model consistency ranged from 92.2% to 95.5%.

Effect of noise

To evaluate the effect of noise on G-causality results, we varied the added noise of each node of the designed structure by a factor ranging from 0 to 0.78 time the standard deviation of the amplitude of each signal. The window length was fixed to 5120 time points, the lags and the connection strength had values as reported in Table 1 for the first and Table 2 for the second example. For the first example, the sensitivity was 1 for the added noise lower or equal to 0.4 (see Figure 7 for 0.2 and 0.4) and it decreased until 0.71, whereas the specificity increased over the noise (see Figure 7 for 0.4 and 0.6 for comparison). The model consistency decreased when the noise increased as a

quadratic function (R^2=0.999) ranging from 97.3% to 93.6% and the model order changed from 21 to 24 time points. The indirect false link was never revealed by G-causality analysis, increasing the noise.

The second example confirmed the results of the first example and it did not provide any additional results.

Discussion

In this study we generated 19 virtual simulated covariance stationary signals with a designed structure of information flows among nodes, which we referred to as a ground truth structure and we performed conditional G-causality analysis in the time domain to test its reliability. Previous simulation studies reported results on the reliability of G-causality concept, but for frequency domain multivariate methods such as Partial Directed Coherence, Directed Transfer Function (DTF), and its modification known as direct DTF (Astolfi et al., 2007), TE and phase slope index (Silfverhuth et al., 2012). In the present study, G-causality measure was applied by means of "Granger Causal Connectivity Analysis" toolbox (Seth, 2010) combined with standard significance testing. The same procedure was applied by a previous simulation study (Haufe et al., 2013) reporting spurious connectivity results regardless of whether the analysis was performed on sensor-space data or on sources estimated using three different established inverse methods. Spurious results of G-causality analyses were attributed to weak data asymmetries caused by linear mixing of the interacting sources, as opposed to strong asymmetries related to genuine time-lagged information flow (Haufe et al., 2013).

In the present study G-causality analysis was performed on simulated data while varying one of the following variables: the number of time points used to perform G-causality, the lags between interacting nodes, the connection strength of the links, and the noise. Based on the ground truth structures, we could estimate the reliability of the G-causality analysis on the identification of the true positive link, on the presence of the false positive link and on the indirect links.

Due to the randomization of the phase of each surrogate signal and to the randomly generated white Gaussian noise, the G-causality results could vary slightly, but these changes did not influenced G-causality reliability and the general conditions which may lead to spurious or missed causalities. In addition the results could change if an alternative less conservative statistic than Bonferroni correction was applied to the coefficients of the MVAR model. The false discovery rate (Benjamini & Hochberg, 1995) could be applied in future studies for comparison.

The main result of this simulation study was that in the presence of covariance stationary signals, window length higher than 1500 time points, model order lower than 40 time points, model consistency higher than 80%, G-causality identified all causal links with the connection strength factor between nodes higher than 0.1 time the amplitude of the driver signal. The number of the false positive links in the G-causality results was more dependent on the length of the temporal window than the lags between causal links, the connection strength and the noise. In the first example if the length of the temporal window was higher than 4000 time points (about 16 s), the G-causality picked all the causal links and the number of false positive links was low (Figure 2, 4352 time points). A previous simulation study with a sampling frequency of 125 Hz reported that 2000 time points (16 s) were sufficient for G-causality analysis (Wang et al., 2014). The number of spurious non-zero G-causality increased when the used number of time points was more than necessary. Indeed if the widow length was higher than 4096 time points the

specificity decreased. This result confirms a previous study claiming the wrong idea of using as much information as possible because the high number of time points could lead to the presence of sampling artifacts in both linear and nonlinear processes (Zhou et al., 2014). The second example did not provide any additional results on the length of the temporal window, as expected, because the number of the signals (19 nodes) and the model order (related to the lags between nodes) are not different between the two examples. Further simulations should be performed with smaller lags to demonstrate that smaller time series would be sufficient for a system of 19 variables.

Another issue in the G-causality analysis is the choice of the model order. An order too low may not allow to describe the data to its full extent, while too big may introduce spurious results (Seth, 2010). For these reasons we evaluated the optimal model order directly by means of the Akaike information criterion. The model order needs to be larger than or at least close to the signal delays (the lags between the causal links), indeed the model order increases when the lags increase. In the first example, when the designed structure had lag values ranging from 1 to 14 time points (the adding value for the lags ranged from -9 to -2) the model order was 23 and all the causal links were detected by G-causality analysis. For the second example, the model order was more close to the maximum lag of the designed structure. Indeed when the maximum lag of the designed structure was 19 the model order was 20. The maximum model order was 35 for lags of the designed structure ranged from 32 to 38 time points. When the lags across nodes were higher than 34 time points (133 ms) or 39 time points (152 ms) for the second example, the sensitivity of G-causality results decreased. This result suggests that G-causality can identify the underlying structure when the lags between signals is lower than about 100 ms.

The variation of the connection strength revealed that G-causality was able to pick all causal links if the connection strength was higher than 0.17 time the amplitude of the driver signal. This result is in accordance with a previous study which showed that when a source influenced a sink with a connection strength lower than 0.2, the causal information was not detected by the G-causality analysis and/or type I errors (the presence of false positive links) were evident (Falasca et al., 2015).

The second example was performed to investigate better the effect of the connection strength on G-causality reliability. In the same designed structure, we simulated links with fixed connection strength and links with varying connection strength factor. Results showed that G-causality did not identify very weak links (i.e. the connection strength factor was equal to 0.05 time the amplitude of the driver signal of the link), whereas it was able to identify links with the connection strength factor higher or equal to 0.1 if the other links of the structure had a connection strength lower than 0.55 time the amplitude of the driver signal. Due to the fact that 0.55 could be considered an high value for interactions in biological systems, this result revealed that G-causality reliability was good also for weak direct causal effects (0.1 time the amplitude of the driver signal).

The simulations showed that the model consistency was affected mainly by the level of the noise. Indeed, the model consistency and the sensitivity decreased when the added uncorrelated noise increased. At high values of noise, G-causality did not pick true causal links, but also false causal links, suggesting that when G-causality analysis reveals a low number of links in real datasets, it is possible that the signal to noise ratio is too low. Previous studies showed that G-causality reliability could be strongly affected by both uncorrelated and linearly mixed additive noise (Nolte et al., 2008; Sommerlade et al., 2012; Friston et al., 2014). In addition uncorrelated noise affected only weakly the

detection of G-causality directionality, whereas linearly mixed noise caused a large fraction of false positives (Vinck et al., 2015).

The evaluation of the G-causality reliability on the indirect link showed that G-causality picked the false link as true when the length of the temporal window was higher than 4864, the lags of the designed structure was lower than 15 time points and the connection strength factor was higher than 0.3 time the amplitude of the driver signal for the first example and 0.5 for the designed structure with higher number of links (second example). To note that in both examples, the connection factor between node 10 to the latent node 20 was fixed to 0.5, whereas the connection factor between the latent node 20 to node 13 varied. Additional simulations should be performed varying also the connection strength factor between node 10 to node 20.

In conclusion, the present simulation study reveals that the number of time points in the temporal window affects G-causality reliability substantially, and rice the issue that a large number of time points could decrease the reliability of the G-causality results, increasing the number of false positive (type I errors). G-causality results are reliable showing all true positive links (absence of type II errors), when the underlying structure has the signal delays between the interacting nodes lower than 100 ms, the connection strength higher to 0.1 time the amplitude of the driver signal and good signal to noise ratio. G-causality detects the indirect link for connection strength higher than the direct link and lags lower than the direct link.

403 **References**

404

405 Akaike H. 1974. A new look at the statistical model identification. *IEEE T. Automat. Contr.* 19:716–723.

406

407 Astolfi L, Cincotti F, Mattia D, Marciani MG, Baccala LA, de Vico Fallani F, Lai M, Baccala L, Salinari S, Ursino
408 M, Zavaglia M, Babiloni F. 2007. Comparison of different cortical connectivity estimators for high-resolution EEG
409 recordings. *Hum. Brain Mapp.* 28:143-157. doi:10.1002/hbm.20263.

410

411 Barnett L, Seth AK. 2013. The MVGC multivariate Granger causality toolbox: a new approach to Granger-causal
412 inference. *J. Neurosci. Methods* 223:50-68. doi: 10.1016/j.jneumeth.2013.10.018.

413

414 Barrett AB, Murphy M, Bruno M-A, Noirhomme Q, Boly M, Laureys S, Seth AK. 2012. Granger Causality
415 Analysis of Steady-State Electroencephalographic Signals during Propofol-Induced Anaesthesia. *PLoS ONE*
416 7:e29072. doi:10.1371/journal.pone.0029072.

417

418 Bastos AM, Schoffelen JM. 2016. A Tutorial Review of Functional Connectivity Analysis Methods and Their
419 Interpretational Pitfalls. *Front. Syst. Neurosci.* 9:175. doi: 10.3389/fnsys.2015.00175.

420

421 Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple
422 testing. J. R. Stat. Soc: Ser B (Methodol.) 57:289–300.

423

424 Bhardwaj R D, Mahmoodabadi SZ, Otsubo H, Carter Snead, OIII, Rutka JT, Widjaja E. 2010. Diffusion tensor
425 tractography detection of functional pathway for the spread of epileptiform activity between temporal lobe and
426 rolandic region. *Childs Nerv. Syst*. 26:185–190. doi:10.1007/s00381-009-1017-1.

427

428 Bressler SL, Seth AK. 2011. Wiener–Granger causality: a well established methodology. *NeuroImage* 58:323–329.
429 doi: 10.1016/j.neuroimage.2010.02.059.

430

431 Brovelli A, Ding MZ, Ledberg A, Chen YH, Nakamura R, Bressler SL. 2004. Beta oscillations in a large-scale
432 sensorimotor cortical network: Directional influences revealed by granger causality. *Proc. Natl. Acad. Sci. U.S.A.*
433 101:9849–9854. doi: 10.1073/pnas.0308538101.

434

435 Coben R, Mohammad-Rezazadeh I. 2015. Neural Connectivity in Epilepsy as Measured by Granger Causality.
436 *Front. Hum. Neurosci.* 9:194. doi: 10.3389/fnhum.2015.00194. Review.

437

438 Cui J, Xu L, Bressler SL, Ding M, Liang H. 2008. BSMART: a Matlab/C toolbox for analysis of multichannel
439 neural time series. *Neural Networks* 21:1094–1104. doi: 10.1016/j.neunet.2008.05.007.

PeerJ Preprints NOT PEER-REVIEWED

Dauwels J, Vialatte F, Musha T, Cichocki A. 2010. A comparative study of synchrony measures for the early diagnosis of Alzheimer's disease based on EEG. *Neuroimage* 49:668-693. doi: 10.1016/j.neuroimage.2009.06.056.

David O, Guillemain I, Saillet S, Reyt S, Deransart C, Segebarth C, Depaulis A. 2008. Identifying neural drivers with functional MRI: an electrophysiological validation. *PLoS Biol.* 6:e315. doi: 10.1371/journal.pbio.0060315.

Ding M, Bressler S, Yang W, Liang H. 2000. Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: data preprocessing, model validation, and variability assessment. *Biol. Cybern.* 83:35–45. doi: 10.1007/s004229900137.

Ding H, Qin W, Liang M, Ming D, Wan B, Li Q, Yu C. 2015. Cross-modal activation of auditory regions during visuo-spatial working memory in early deafness. *Brain* 138:2750-2765. doi: 10.1093/brain/awv165.

Falasca NW, D'Ascenzo S, Di Domenico A, Onofrj M, Tommasi L, Laeng B, Franciotti R. 2015. Hemispheric lateralization in top-down attention during spatial relation processing: a Granger causal model approach. *Eur. J. Neurosci* 41:914-924. doi: 10.1111/ejn.12846.

Franciotti R, Falasca NW, Bonanni L, Anzellotti F, Maruotti V, Comani S, Thomas A, Tartaro A, Taylor JP, Onofrj M. 2013. Default network is not hypoactive in dementia with fluctuating cognition: an Alzheimer disease/dementia with Lewy bodies comparison. *Neurobiol. Aging* 34:1148-1158. doi: 10.1016/j.neurobiolaging.2012.09.015.

Fries P. 2015. Rhythms for cognition: communication through coherence. *Neuron* 88:220–235. doi: 10.1016/j.neuron.2015.09.034.

Friston KJ. 2011. Functional and effective connectivity: a review. *Brain Connect*. 1:13-36. doi: 10.1089/brain.2011.0008. Review.

Friston KJ, Bastos AM, Oswal A, van Wijk B, Richter C, Litvak V. 2014. Granger causality revisited. *NeuroImage* 101:796-808. doi: 10.1016/j.neuroimage.2014.06.062.

Gao Q, Duan X, Chen H. 2011. Evaluation of effective connectivity of motor areas during motor imagery and execution using conditional Granger causality. *Neuroimage* 54:1280-1288. doi: 10.1016/j.neuroimage.2010.08.071.

Geweke J. 1982. Measurement of linear dependence and feedback between multiple time series. *J. Am. Stat. Assoc,* 77:304–313. doi: 10.2307/2287238.

Granger CW. 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econ. J. Econ. Soc.* 37:424–438. doi: 10.2307/1912791.

Granger C, Newbold P. 1974. Spurious regressions in econometrics. *J. Econom* 2:111–120. doi: 10.1136/jech-2013-202645.

Haufe S, Nikulin VV, Müller KR, Nolte G. 2013. A critical assessment of connectivity measures for EEG data: a simulation study. *Neuroimage* 64:120-133. doi: 10.1016/j.neuroimage.2012.09.036.

Hesse W, Moller E, Arnold M, Schack B. 2003. The use of time-variant EEG granger causality for inspecting directed interdependencies of neural assemblies. *J. Neurosci. Methods* 124:27–44.

Nicolaou N, Hourris S, Alexandrou P, Georgiou J. 2012. EEG-Based Automatic Classification of 'Awake' versus 'Anesthetized' State in General Anesthesia Using Granger Causality. *PLoS ONE* 7:e33869. doi:10.1371/journal.pone.0033869.

Nolte G, Ziehe A, Nikulin VV, Schlögl A, Krämer N, Brismar T, Müller KR. 2008. Robustly estimating the flow direction of information in complex physical systems. *Phys. Rev. Lett.* 100:234101. doi:10.1103/PhysRevLett.100.234101.

Protopapa F, Siettos CI, Evdokimidis I, Smyrnis N. 2014. Granger causality analysis reveals distinct spatio-temporal connectivity patterns in motor and perceptual visuo-spatial working memory. *Front. Comput. Neurosci.* 8:146. doi: 10.3389/fncom.2014.00146.

Protopapa F, Siettos CI, Myatchin I, Lagae L. 2016. Children with well controlled epilepsy possess different spatio-temporal patterns of causal network connectivity during a visual working memory task. *Cogn Neurodyn*. 10:99–111. doi: 10.1007/s11571-015-9373-x.

Seth AK. 2010. A MATLAB toolbox for Granger causal connectivity analysis. *J. Neurosci. Methods* 186:262–273. doi: 10.1016/j.jneumeth.2009.11.020.

Seth AK, Barrett AB, Barnett L. 2015. Granger causality analysis in neuroscience and neuroimaging. *J. Neurosci.* 35:3293-3287. doi: 10.1523/JNEUROSCI.4399-14.2015.

Siggiridou E, Kugiumtzis D. 2016. Granger causality in multivariate time series using a time-ordered restricted vector autoregressive model. IEEE Trans. Signal Process. 64:1759-1773 .

514 Silfverhuth MJ, Hintsala H, Kortelainen J, Seppänen T. 2012. Experimental comparison of connectivity measures
515 with simulated EEG signals. *Med. Biol. Eng. Comput.* 50:683-638. doi: 10.1007/s11517-012-0911-y.
516
517 Sommerlade L, Thiel M, Platt B, Plano A, Riedel G, Grebogi C, Timmer J, Schelter B. 2012. Inference of granger
518 causal time-dependent influences in noisy multivariate time series. J*ournal of neuroscience methods* 203:173–185.
519 doi: 10.1016/j.jneumeth.2011.08.042.
520
521 Varela F, Lachaux JP, Rodriguez E, Martinerie J. 2001. The brainweb: phase synchronization and large-scale
522 integration. *Nat. Rev. Neurosci.* 2:229–239. doi: 10.1038/35067550.
523
524 Vinck M, Huurdeman L, Bosman CA, Fries P, Battaglia FP, Pennartz CM., Tiesinga PH. 2015. How to detect the
525 Granger-causal flow direction in the presence of additive noise? *Neuroimage* 108:301-318. doi:
526 10.1016/j.neuroimage.2014.12.017.
527
528 Vlachos I, Kugiumtzis D. 2013. Backward-in-time selection of the order of dynamic regression prediction model.
529 Journal of Forecasting 32:685–701.
530
531 Wiener N. 1956. “The theory of prediction,” in Modern Mathematics for Engineers, ed E. Beckenbach (New York,
532 NY: McGraw-Hill), 165–190.
533
534 Wang HE, Bénar CG, Quilichini PP, Friston KJ, Jirsa VK., Bernard C. 2014. A systematic framework for functional
535 connectivity measures. *Front. Neurosci.* 8:405. doi: 10.3389/fnins.2014.00405.
536
537 Zhou D, Zhang Y, Xiao Y, Cai D. 2014. Reliability of the Granger causality inference. *New J. Phys.* 16:043016.

PeerJ Preprints NOT PEER-REVIEWED

Figure legend

Figure 1 Schematic representation of the designed structures for the two examples. The links of the designed structures are shown by the G-causality matrices with their connection strength factors (as indicated by the colorbars) for the first and the second examples (A), and by arrows indicating the direction and versus of the links between nodes (B).

Figure 2 G-causality results of the first example for different values of the window lengths. (A) The G-causality matrices show the G-magnitude values (as indicated by the colorbars) for the causal links at 1792, 4352 and 7424 time points. Pink circles indicate the true positive links. (B) Graphs of the results. The thickness of the arrows indicates the values of the G-magnitude for the identified links.

Figure 3 G-causality results of the first example for different values of lags. (A) The G-causality matrices show the G-magnitude values (as indicated by the colorbars) for the causal links for adding values of the lags of the starting structure of -9, 5 and 29 time points. Pink circles indicate the true positive links. (B) Graphs of the results. The thickness of the arrows indicates the values of the G-magnitude for the identified links.

Figure 4 Graphs of the results of the second example while varying lag values between links. Results show the identified links by G-causality for a designed structure with lag ranging from 3 to 9 time points (adding value to the starting structure was -7), ranging from 10 to 16 time points (no adding value), ranging from 39 to 45 time points (adding value of 29).

Figure 5 G-causality results of the first example for different values of the connection strength. (A) The G-causality matrices show the G-magnitude values (as indicated by the colorbars) for adding values of the connection strength factor of -0.2, 0.05 and 0.1. Pink circles indicate the true positive links. (B) Graphs of the results. The thickness of the arrows indicates the values of the G-magnitude for the identified links.

Figure 6 Graphs of the results of the second example while varying connection strength factor between links. Numbers indicate the adding value to the connection strength factors of the starting designed structure of the Table 2.

Figure 7 G-causality results of the first example for different values of noise. (A) The G-causality matrices show the G-magnitude values (as indicated by the colorbars) for adding value of the noise of 0.2, 0.4 and 0.6 the standard deviation of the amplitude of each signal. Pink circles indicate the true positive links. (B) Graphs of the results. The thickness of the arrows indicates the values of the G-magnitude for the identified links.

PeerJ Preprints NOT PEER-REVIEWED

Table 1(on next page)

First example.

Starting values of the lags of the causal links and the factors which multiply the amplitude of the driver signal of the link defining the connection strength of the designed structure.

PeerJ Preprints NOT PEER-REVIEWED

Table 1 First example. Starting values of the lags of the causal links and the factors which multiply the amplitude of the driver signal of the link defining the connection strength of the designed structure.

Links	Lags (time points)	Connection strength factor
from node 2 to node 4	10	0.4
from node 4 to node 2	12	0.5
from node 4 to node 15	13	0.45
from node 8 to node 6	14	0.3
from node 11 to node 9	16	0.35
from node 11 to node 18	14	0.3
from node 17 to node 7	15	0.55

PeerJ Preprints NOT PEER-REVIEWED

Table 2(on next page)

Second example.

Starting values of the lags of the links and the connection strength factors which multiply the amplitude of the driver signal of the link. An “*” was added to the connection strength factor which changed during the simulation for the study of the influence of connection strength to G-causality results.

PeerJ Preprints NOT PEER-REVIEWED

Table 2 Second example. Starting values of the lags of the links and the connection strength factors which multiply the amplitude of the driver signal of the link. An "*" was added to the connection strength factor which changed during the simulation for the study of the influence of connection strength to G-causality results.

Links	Lags (time points)	Connection strength factor
from node 1 to node 3	14	0.2 *
from node 3 to node 6	11	0.05
from node 5 to node 8	10	0.2
from node 10 to node 9	12	0.25
from node 13 to node 5	13	0.05
from node 7 to node 19	16	0.1
from node 19 to node 3	15	0.15
from node 2 to node 4	10	0.4 *
from node 4 to node 2	12	0.5 *
from node 4 to node 15	13	0.45 *
from node 8 to node 6	14	0.3 *
from node 11 to node 9	16	0.35 *
from node 11 to node 18	14	0.3 *
from node 17 to node 7	15	0.55 *

PeerJ Preprints NOT PEER-REVIEWED

Figure 1(on next page)

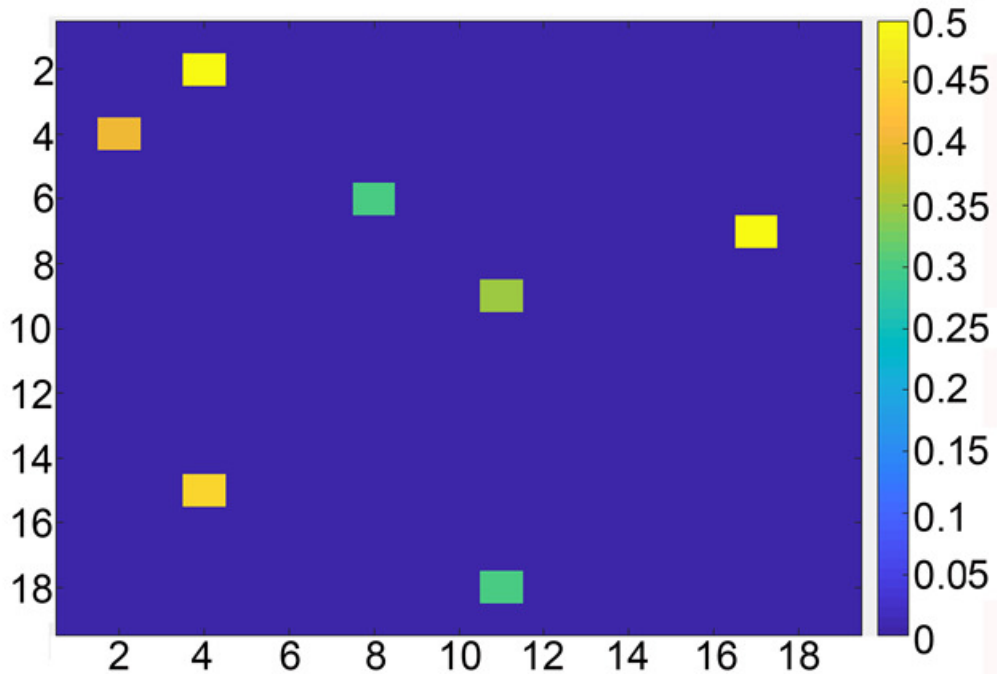
Schematic representation of the designed structures for the two examples.

The links of the designed structures are shown by the G-causality matrices with their connection strength factors (as indicated by the colorbars) for the first and the second examples (A), and by arrows indicating the direction and versus of the links between nodes (B).

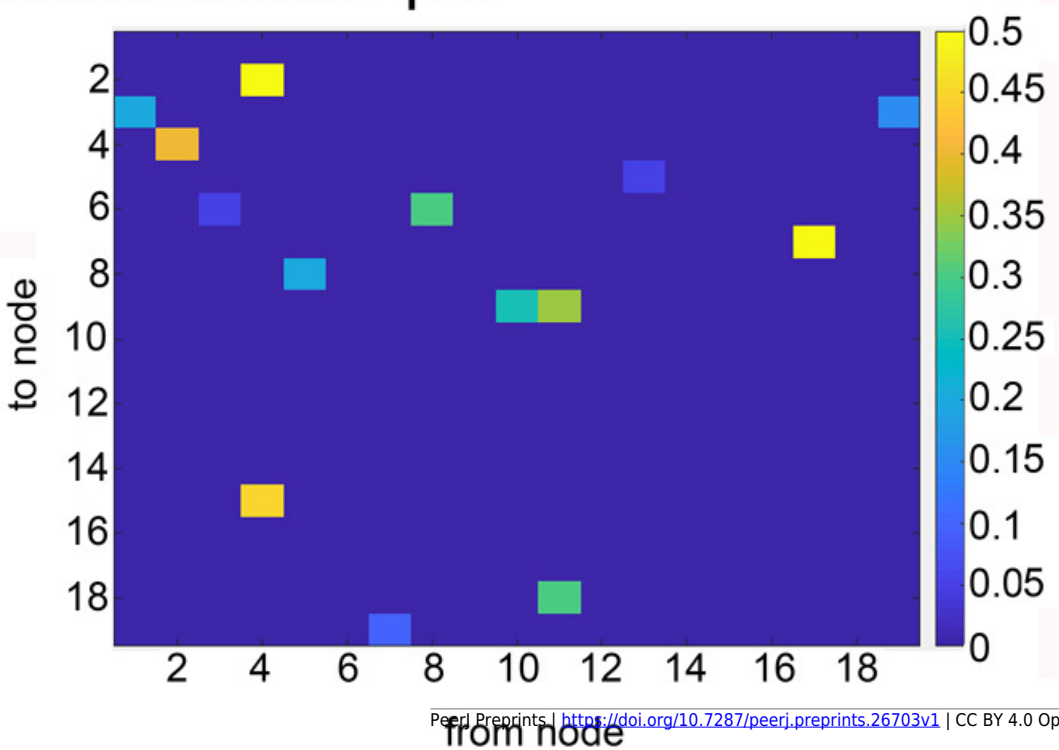
PeerJ Preprints NOT PEER-REVIEWED

A

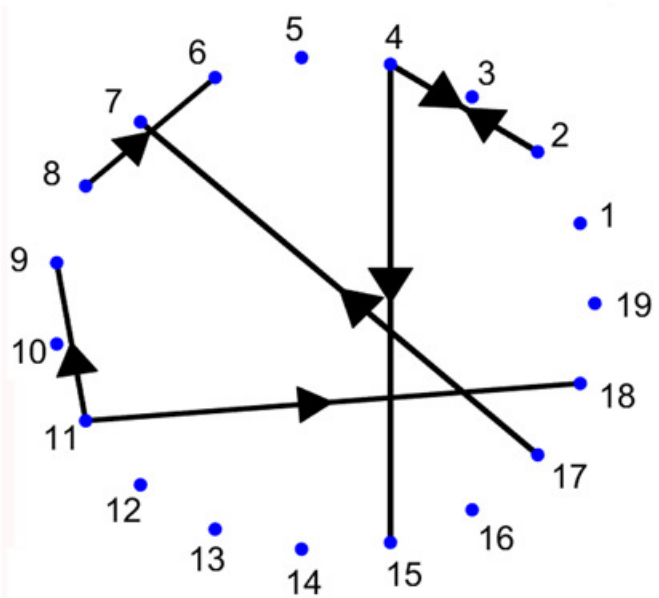
First example

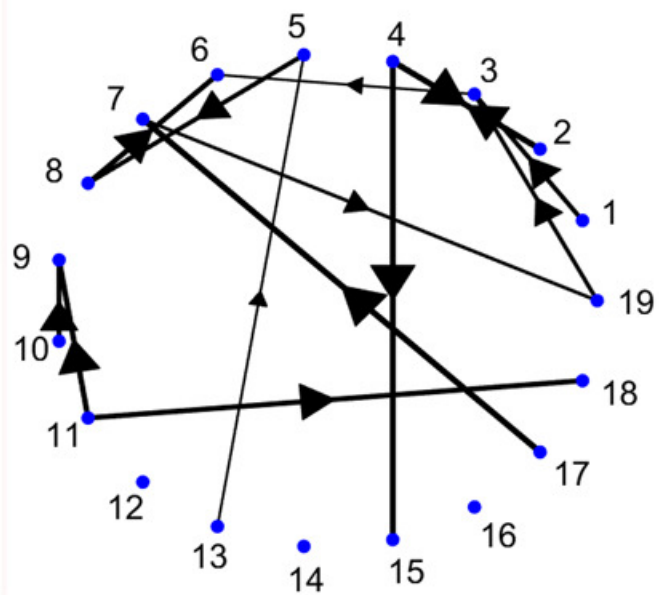


Second example



B





PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26703v1 | CC BY 4.0 Open Access | rec: 15 Mar 2018, publ: 15 Mar 2018

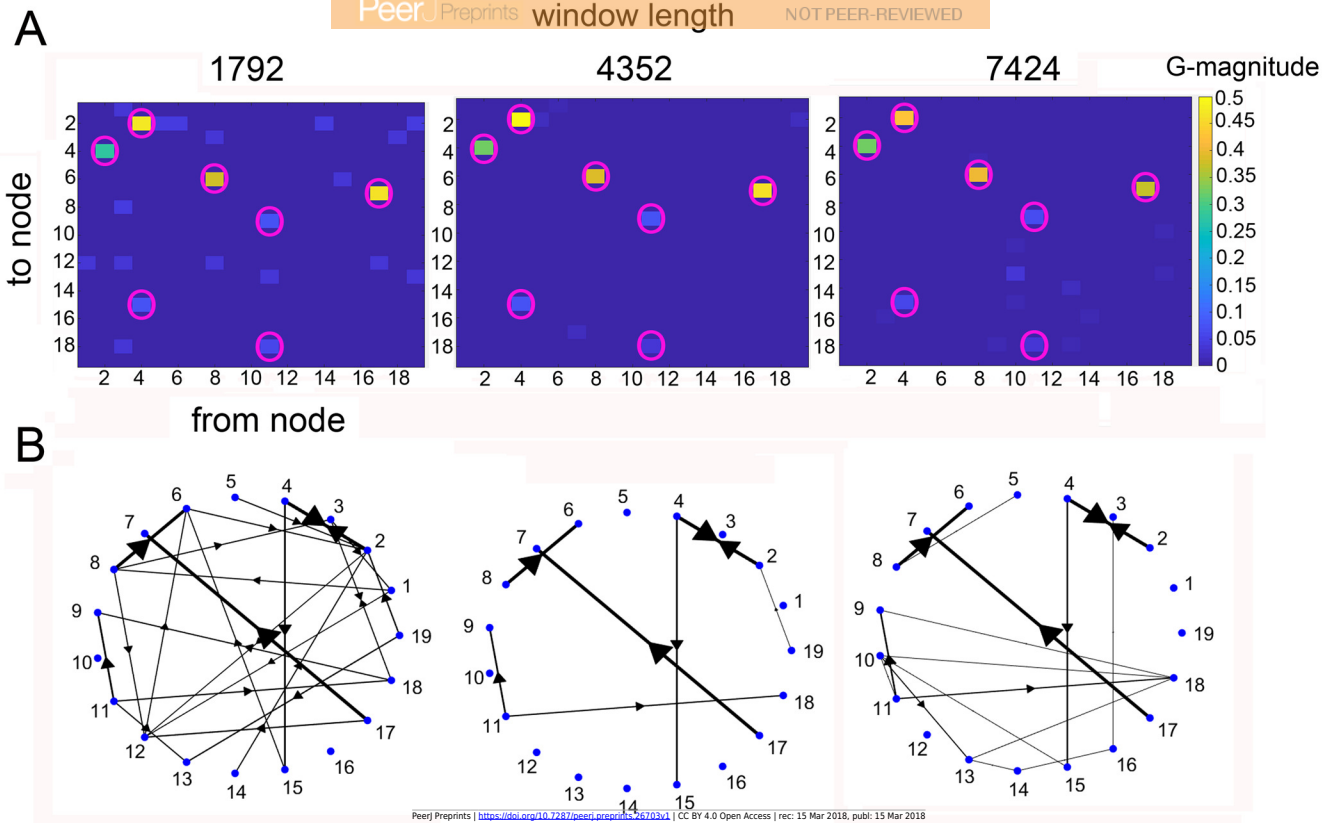
PeerJ Preprints NOT PEER-REVIEWED

Figure 2(on next page)

G-causality results of the first example for different values of the window lengths.

(A) The G-causality matrices show the G-magnitude values (as indicated by the colorbars) for the causal links at 1792, 4352 and 7424 time points. Pink circles indicate the true positive links. (B) Graphs of the results. The thickness of the arrows indicates the values of the G-magnitude for the identified links.

PeerJ Preprints window length NOT PEER-REVIEWED



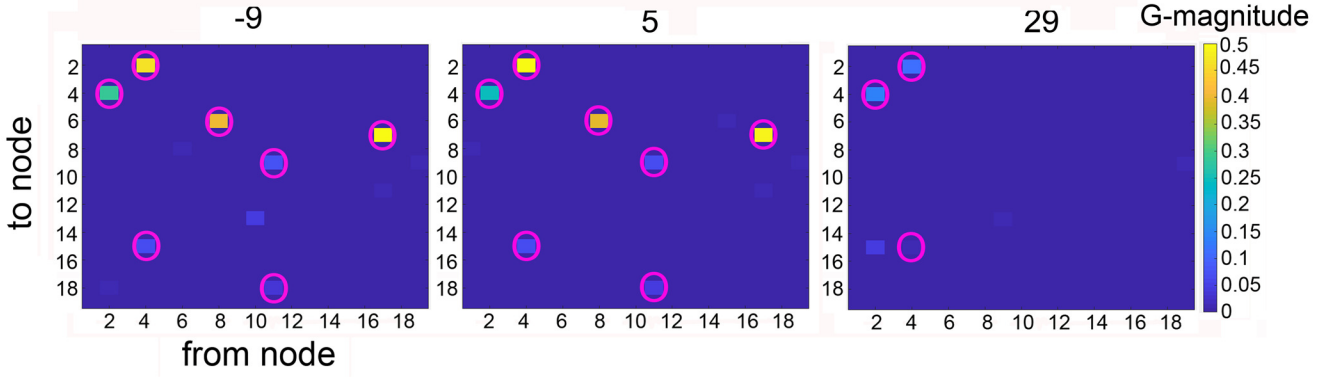
PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26703v1 | CC BY 4.0 Open Access | rec: 15 Mar 2018, publ: 15 Mar 2018

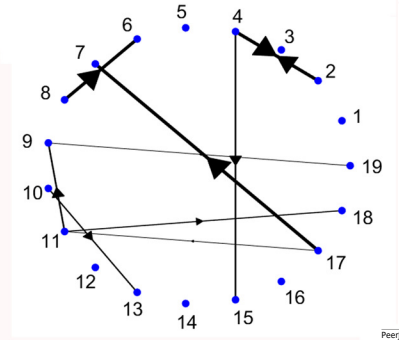
Figure 3(on next page)

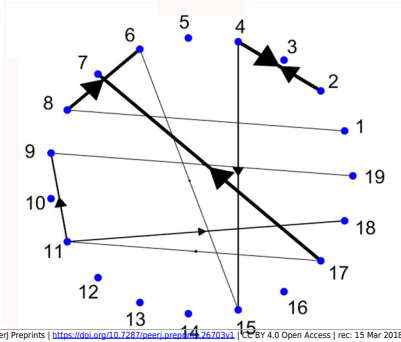
G-causality results of the first example for different values of lags.

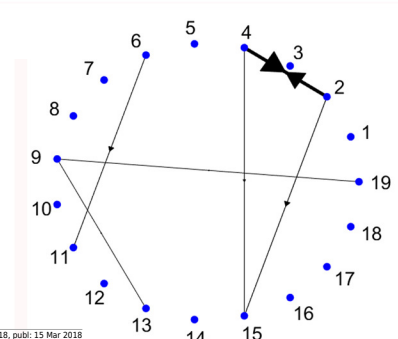
(A) The G-causality matrices show the G-magnitude values (as indicated by the colorbars) for the causal links for adding values of the lags of the starting structure of -9, 5 and 29 time points. Pink circles indicate the true positive links. (B) Graphs of the results. The thickness of the arrows indicates the values of the G-magnitude for the identified links.

PeerJ Preprints









NOT PEER-REVIEWED

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26703v1 | CC BY 4.0 Open Access | rec: 15 Mar 2018, publ: 15 Mar 2018

PeerJ Preprints NOT PEER-REVIEWED

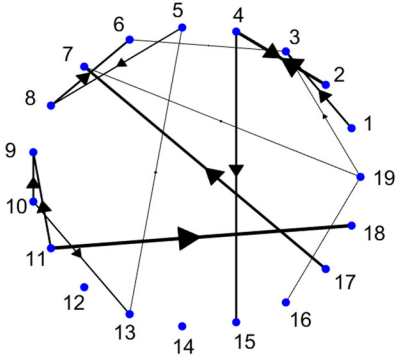
Figure 4(on next page)

Graphs of the results of the second example while varying lag values between links.

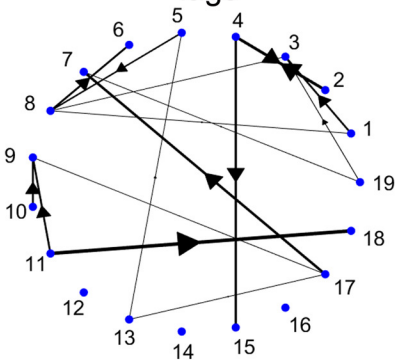
Results show the identified links by G-causality for a designed structure with lag ranging from 3 to 9 time points (adding value to the starting structure was -7), ranging from 10 to 16 time points (no adding value), ranging from 39 to 45 time points (adding value of 29).

Peer Preprints NOT PEER-REVIEWED

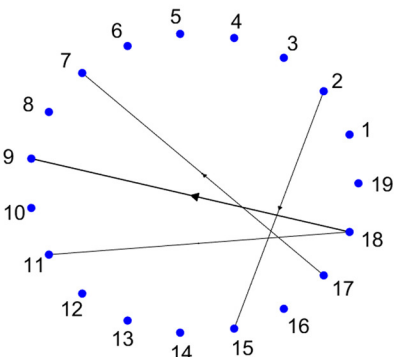
lags



-7



0



29

Peer Preprints | https://www.preprints.org/manuscript/201803.0144/v1 © 2018 Open Access | rec: 15 Mar 2018, publ: 15 Mar 2018

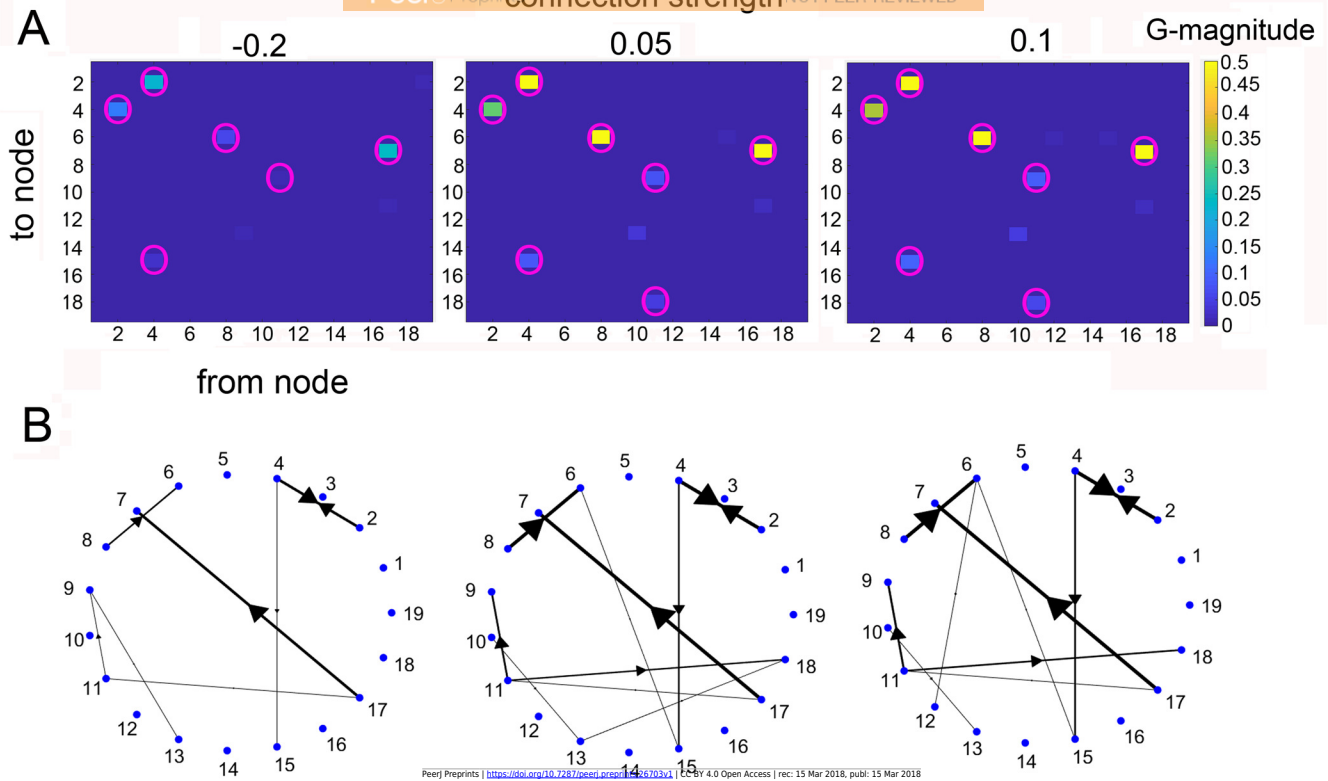
PeerJ Preprints NOT PEER-REVIEWED

Figure 5(on next page)

G-causality results of the first example for different values of the connection strength.

(A) The G-causality matrices show the G-magnitude values (as indicated by the colorbars) for adding values of the connection strength factor of -0.2, 0.05 and 0.1. Pink circles indicate the true positive links. (B) Graphs of the results. The thickness of the arrows indicates the values of the G-magnitude for the identified links.

PeerJ Preprints connection strength NOT PEER-REVIEWED



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26703v1 | CC BY 4.0 Open Access | rec: 15 Mar 2018, publ: 15 Mar 2018

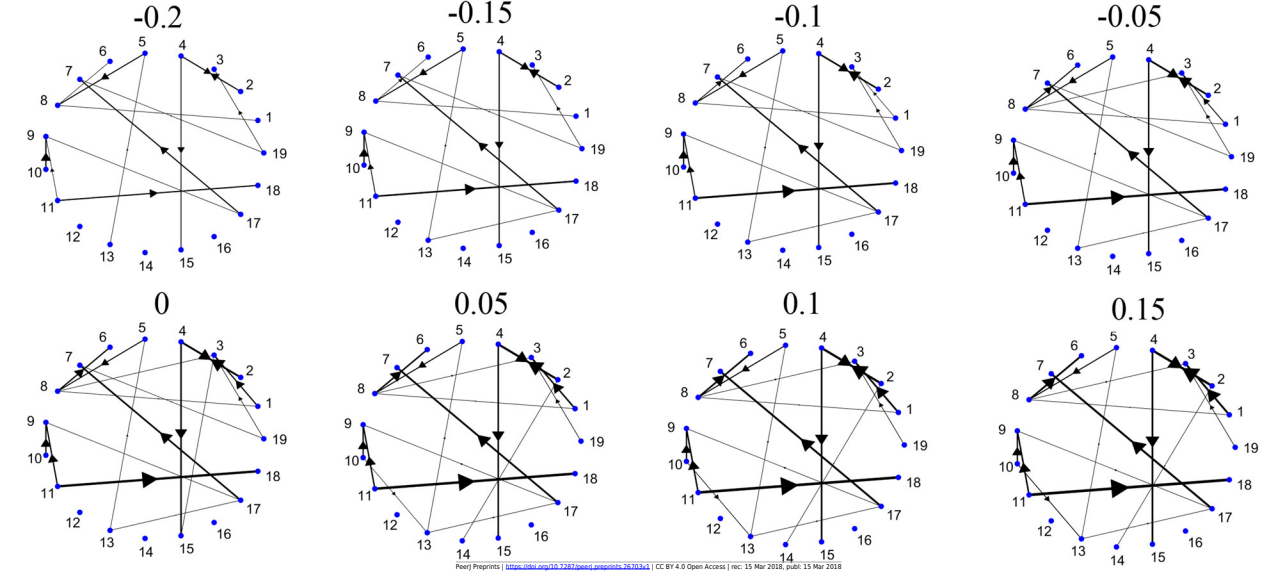
Figure 6(on next page)

Graphs of the results of the second example while varying connection strength factor between links.

Numbers indicate the adding value to the connection strength factors of the starting designed structure of the Table 2.

PeerJ Preprints | PEER-REVIEWED

connection strength



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26703v1 | CC BY 4.0 Open Access | rec: 15 Mar 2018, publ: 15 Mar 2018

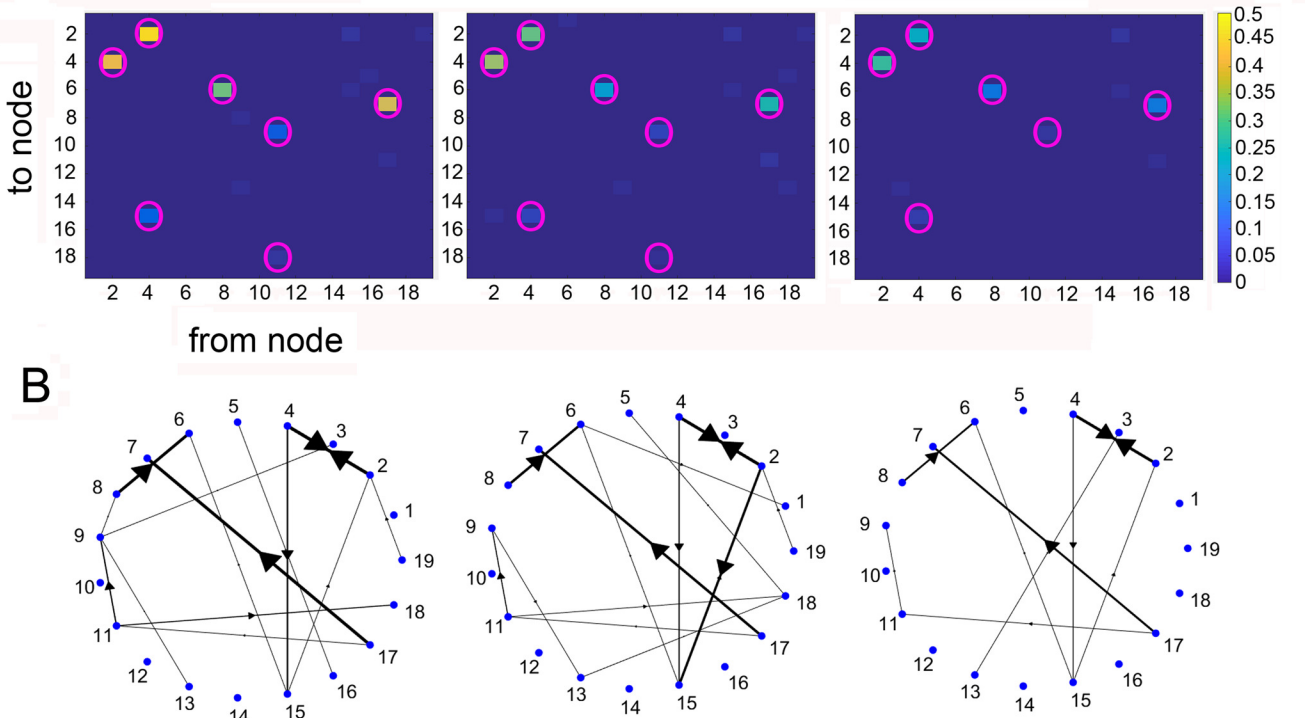
PeerJ Preprints NOT PEER-REVIEWED

Figure 7 (on next page)

G-causality results of the first example for different values of noise.

(A) The G-causality matrices show the G-magnitude values (as indicated by the colorbars) for adding value of the noise of 0.2, 0.4 and 0.6 the standard deviation of the amplitude of each signal. Pink circles indicate the true positive links. (B) Graphs of the results. The thickness of the arrows indicates the values of the G-magnitude for the identified links.

PeerJ Preprints noise NOT PEER-REVIEWED



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26703v1 | CC BY 4.0 Open Access | rec: 15 Mar 2018, publ: 15 Mar 2018