Beyond element-wise interactions: defining group-to-group interactions for biological processes

Christophe Ladroue[1] Shuixia Guo[2] Keith Kendrick[3] Jianfeng Feng[1,4*]

Department of Computer Science and Mathematics, Warwick University, Coventry CV4 7AL, UK

[2]Mathematics and Computer Science College, Hunan Normal University, Changsha 410081, P.R.China

[3]Laboratory of Behaviour and Cognitive Neuroscience, The Babraham Institute, Cambridge, UK

[4]Centre for Computational System Biology, Fudan University, PR China

* Corresponding author: Jianfeng.Feng@warwick.ac.uk

Abstract

Background We are often faced with the situation that genes, proteins and molecules work cooperatively or competitively to achieve a task when we study the properties of their interacting networks. It is of substantial importance to study the interactions among *groups* of nodes, which we call complex interactions. **Methodology** To tackle this problem, we extend traditional Granger causality to *complex Granger causality*, both in time domain and in frequency domain. Our method is extensively tested in synthesized data and three biological datasets: microarray data of the *Saccharomyces cerevisiae* yeast, local field potential recordings of 2 brain areas and a metabolic reaction. **Conclusions** Our results demonstrate that *complex Granger causality* can reveal the intrinsical stucture of the networks and is more useful than traditional Granger causality. Our approach raises some fundamental issues of the systems biology approach.

Author summary

Uncovering the existence and direction of interactions between elements of a set of signals remains a difficult and harduous task that one has to face if one wants to understand the mechanisms at work in most of biological phenomena and make full use of the high throughput experimental data that is now more and more available. The picture appears even more daunting when the possibility of group-wise interactions is taken into account. We propose an extension of Granger causality which allows the quantification of the strength of connectivity between groups of signals, and not only between individual signals. This new measure is defined in the frequency domain as well as the time domain and can reveal complex interactions of elements working cooperatively or competitively, giving a more complete view of the mechanism in place. It is also further extended to partial complex Granger causality, in order to minimize the influence of other signals. We apply our method on biological problems and show how grouping signals together greatly improves the information we can obtain from the data, as well as possibly reveals connections that would be missed with an element-wise method.

1 Introduction

A network structure carefully inferred from experimental data could provide us with critical information about the underlying system of investigation and is an important topic in system biology. For example, high-throughput data from gene, metabolic, signaling or transcriptional regulatory networks, contain information of thousands of genes or products thereof. For such complex networks, complex (group) interactions occur since the nodes (genes, proteins and substances) may work cooperatively or competitively to accomplish a task. The complex interactions are considerably different from the interactions among single node which have been extensively studied in the past decades. For example, there could be a situation where both of two single nodes have no interaction to a third node, but when they combine together, they together interact with the third one. On the other hand, when two nodes have a negative correlation and each of them interacts with the third one, the interaction may disappear when they are grouped together. Another example is to consider the chemical reaction from S to P. An enzyme acts here as a catalyst from S to P, but not from P to S. Hence to find from observed data that E is a cause of the rate from S to P, but not from P to S is obviously an interesting and challenging task (see Fig. 1). To fully understand the properties of a network, whether it is a gene, a substance, a protein or a neuronal network, it is therefore of prominent importance to consider complex interactions.

Experimentally this issue has been realized and tested intensively. For example, LOF (loss of function) experiments are performed for double, triple and quadruple mutations. Two commonly used computational approaches to explore the experimental data and recover the interactions between units in the literature are Bayesian network [1] and Granger causality [2, 3, 4, 5, 6]. However, to the best of our knowledge, no systematic approach has been developed to take this issue into account. Here we adopt the Granger causality approach. The concept of the Granger causality originally introduced by Wiener [7] and formulated by Granger [2] has played a considerably important role in investigating the re-

lationship among stationary time series. Specifically, given two time series, if the variance of the prediction error for the second time series at the present time is reduced by including past measurements from the first time series in the (non)linear regression model, then the first time series can be said to cause the second time series. Geweke's decomposition of a vector autoregressive process [8, 9, 10, 11] led to a set of causality measures which have a spectral representation and make the interpretation more informative and useful.

To tackle the issue of complex interactions, in the current paper, we extend the pairwise Granger causality and the partial Granger causality we proposed in []12,13] to *complex Granger causality*, both in the time and frequency domains. The previous methods were limited to the study of interactions with a single signal. We apply our approach to synthesized and experimental data to validate the efficiency of our approach. In synthesized data, we first demonstrate that our *complex Granger causality* can reliably detect the pairwise complex interactions, both in the time and frequency domains. Purely in terms of the data, complex interactions could arise even when there are no interactions between single nodes. In order to deal with the interactions between multi-complexes, we then extend the *pairwise complex Granger causality* to *partial complex Granger causality*.

Complex Granger causality is then applied to three different biological problems in order to illustrate its ability to capture these new types of interactions (group-to-signal, group-to-group and group-to-interaction). First, we use yeast cell-cycle microarray data to compare results obtained when complexes-to-single gene connections are not taken into account and when they are. Next, we use complex Granger causality to study the connections between brain areas and compare the results obtained from considering individual signals alone or region averages. Last, we consider a well-known metabolic reaction and show that our method can capture the effect of an enzyme over a chemical reaction rate.

2 Results

2.1 Synthesized data: pairwise complex interaction

Example 1 Suppose that 2 simultaneously generated multiple time series are defined by the equations

$$\left\{ \begin{array}{rcl} \overrightarrow{X_t} & = & A_{11}\overrightarrow{X}_{t-1} + A_{12}\overrightarrow{X}_{t-2} + \overrightarrow{\epsilon_{1t}} \\ \overrightarrow{Y_t} & = & A_{21}\overrightarrow{X}_{t-1} + A_{22}\overrightarrow{X}_{t-2} + B_{21}\overrightarrow{Y}_{t-1} + \overrightarrow{\epsilon_{2t}} \end{array} \right.$$

where $\overrightarrow{X_t} = (x_{1t}, x_{2t})^T$ is a 2-dimensional vector, $\overrightarrow{Y_t} = (x_{3t}, x_{4t}, x_{5t})^T$ is a 3-dimensional vector, $\overrightarrow{\epsilon_{1t}}, \overrightarrow{\epsilon_{2t}}$ are normally distributed random vectors. The coefficient matrices are

$$A_{11} = \begin{pmatrix} 0.95\sqrt{2} & 0 \\ 0 & 0 \end{pmatrix}, A_{12} = \begin{pmatrix} -0.9025 & 0 \\ -0.5 & 0 \end{pmatrix}, A_{21} = \begin{pmatrix} 0.1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, A_{22} = \begin{pmatrix} 0 & 0.4 \\ 0 & 0 \\ 0 & 0 \end{pmatrix},$$

$$B_{21} = \begin{pmatrix} 0 & 0 & 0 \\ -0.5 & 0.25\sqrt{2} & 0.25\sqrt{2} \\ 0 & -0.25\sqrt{2} & 0.25\sqrt{2} \end{pmatrix}$$

We perform a simulation of this system to generate a dataset of 2000 data points with a sample rate of 200 Hz. The trace of the two vectors $\overrightarrow{X_t}$ and $\overrightarrow{Y_t}$ are plotted in Fig. 2 (**A**) (inside ovals). From the model, $\overrightarrow{X_t}$ is clearly a direct source of $\overrightarrow{Y_t}$, which in turn does not have any influence on $\overrightarrow{X_t}$, as represented in Fig. 2.

Fig 2 (**B**) presents a comparison between the time domain pairwise complex Granger causality and the frequency domain pairwise complex Granger causality (see Fig 2 (**C**) for details). Blue lines are the value of the pairwise complex Granger causality calculated in the time domain. Red dots are the summation (integration) of the pairwise complex Granger causality for frequencies in the range of $[-\pi, \pi]$. Fig 2 (**C**) is the results obtained in the frequency domain, as expected, Fig 2 (**C**) demonstrates that the decomposition in the frequency domain fits very well with the pairwise Granger causality in the time domain. The direct causal link from $\overrightarrow{X_t}$ to $\overrightarrow{Y_t}$ is clearly seen, as well as the absence of interaction from $\overrightarrow{Y_t}$ to $\overrightarrow{X_t}$, as expected.

This example clearly demonstrates that the *complex Granger causality* can detect interactions between complexes.

In the next example, we show how a group of signals can have a significant impact on another signal, even if the individual interactions are too small to be detected.

Example 2 Consider the following model

$$\begin{cases} X_1(t) &= 0.8X_1(t-1) - 0.5X_1(t-2) + aX_4(t-1) + \epsilon_1(t) \\ X_2(t) &= 0.6X_2(t-1) - 0.4X_2(t-2) + \epsilon_2(t) \\ X_3(t) &= 0.9025X_3(t-1) - 0.7X_3(t-2) + \epsilon_3(t) \\ X_4(t) &= 0.3X_2(t) - 0.25X_3(t) + \epsilon_4(t); \end{cases}$$

where a is a constant, $\epsilon_i, i = 1, 2, 3, 4$ are independent standard normal random variables. The time traces of $X_i(t), i = 1, 2, 3, 4$ with $a = 0.3$ are shown in Fig. 3 (**A**). The parameter a allows us to control how much influence a combination of X_2 and X_3 has on X_1.

In Fig. 3B, the mean values of the Granger causality together with their 3σ confidence intervals are depicted (after 1000 replications). Treating X_2 and X_3 as a single units shows no interaction to X_1 (the lower bounds of the confidence intervals of $F_{3\to1}$ and $F_{2\to1}$ are both less than 0). However their combination as X_4 shows a significant interaction with X_1.

In Fig. 3C, we plot the lower bound of the confidence interval of Granger causalities as a function of a. We see that the connection $X_4 \to X_1$ is much more significant than $X_2 \to X_1$ and $X_3 \to X_1$, and is so much more rapidly. Replacing X_4 by its expression in X_1 in example 2, we see that the contributions of X_2 and X_3 are respectively $0.3a$ and $0.25a$. These remain small for small values of a, hence the low value of their corresponding interactions with X_1. As a increases however, these two contributions become larger and thus their Granger causalities as well, as seen in Fig 3C, inset.

2.2 Synthesized data: partial complex interaction

Indirect connections can produce spurious links between groups of interest. We have extended the method further with *partial complex Granger causality*, which estimates the complex Granger causality while reducing the influence of a third group.

Example 3 We modify example 1 to the following model

$$\begin{cases} \overrightarrow{X_t} &= A_{11}\overrightarrow{X}_{t-1} + A_{12}\overrightarrow{X}_{t-2} + \overrightarrow{\epsilon_{1t}} \\ \overrightarrow{Y_t} &= A_{21}\overrightarrow{X}_{t-1} + A_{22}\overrightarrow{X}_{t-2} + B_{21}\overrightarrow{Y}_{t-1} + \overrightarrow{\epsilon_{2t}} \\ \overrightarrow{Z_t} &= B_{31}\overrightarrow{Y}_{t-1} + C_{31}\overrightarrow{Z}_{t-1} + \overrightarrow{\epsilon_{3t}} \end{cases}$$

where $\overrightarrow{Z_t} = (x_{6t}, x_{7t})^T$, $B_{31} = \begin{pmatrix} -0.95 & 0 & 0.1 \\ -0.5 & 0 & 0 \end{pmatrix}$, $C_{31} = \begin{pmatrix} 0.7 & 0.2 \\ 0.6 & 0 \end{pmatrix}$

We perform a simulation of this system to generate a data set of 2000 data points with a sample rate of 200 Hz. Traces of the time series are plotted in Fig 4 (**A**), both X, Y and Z are complexes (in circles). From the model, we see that only $X \rightarrow Y$ and $Y \rightarrow Z$ are direct interactions, as depicted in Fig 4 (**A**). Partial complex Granger causality and its 95% confidence intervals after 1000 replications are calculated. Fig 4 (**B**) presents a comparison between the time domain partial complex Granger causality and the frequency domain partial complex Granger causality. The blue solid line is the partial complex Granger causality and its 95% confidence intervals while the red dotted line is the integration of the partial complex Granger causality for frequencies in the range of $[-\pi, \pi]$. Fig 4 (**C**) shows the results obtained in the frequency domain. They are in very good agreement with those in the time domain.

Applying a pairwise complex Granger causality from X to Z gives a value of 0.2, which is misleadingly high given the indirect nature of their connection. In contrast, considering the partial complex Granger causality $X \rightarrow Z|Y$ removes the influence of Y and gives a more accurate value of 0: Z can be completely explained in terms of Y alone.

2.3 Complexes in the yeast cell-cycle

We now apply our method to the binding interactions of proteins during the cell cycle of the budding yeast *Saccharomyces cerevisiae*. A gene can be activated by a combination of multiple transcription factors (a complex) and our aim is to show that grouping those transcription factors that act together strengthens the connection to their target genes. We use the microarray data produced for a study of the yeast's cell cycle ([12, 13]). We selected 12 time traces corresponding to either transcription factors or cell-cycle related genes. Among the 5 transcriptions factors, we know that some belong to the same complexes (MBP1 and SWI6, SWI4 and SWI6, from MIPS, [14]) and we expect their combination to have a stronger effect than when considered individually.

In order to test this claim, we apply pairwise Granger causality on all pairs (Transcription Factor, Gene) and (Complex, Gene). The inferred network is compared against the true network, built from the up-to-date data available on the curated YEASTRACT database ([15]). The resulting network is shown on figure 5. The program missed only one interaction (thin dashed line) and most of the calculated connections are true positives

(thick lines) - they do exist in the true network, either from documented evidence (blue lines) or marked as potential (green lines) in YEASTRACT. The thin blue lines represent false positives, *i.e.* links suggested by the causality measure but not found in the literature. Most of the network is very close to the true network.

As seen on figure 5B, using a complex can greatly improve the strength of the interactions between transcription factors and target genes. Connections that could have been erroneously discarded at low threshold are more likely to be kept once the complex is considered.

2.4 Change in connections of brain regions during face recognition

In neuroscience, it is often of great importance to uncover connections between brain regions. Since most techniques are based on the interactions between individual signals, a workaround is usually to average them over a region of interest (see [17] for an example on fMRI data) beforehand. This can be misleading as a (weighted) average cannot capture the overall effect of individual interactions. It is especially true when spatial resolution is very high: interactions between groups of neurons are much more informative than those between individual neurons. In this section, we consider the neuronal activity of the left and right inferior temporal cortex (IT) in a sheep's brain, before and during a visual stimulus. The recording was carried when the sheep looked at a fixation point for one second and then an image (a pair of faces) for one second. We have the recordings of 64 local field potentials in each region, sampled at a rate of 1000 Hz. The animal was handled in strict accordance with good animal practice as defined by the UK Home Office guidelines, and all animal work was approved by the appropriate committee.

We select the time series with significant variation (standard deviation > 0.01). After this filtering, the left and right regions contain respectively 10 and 11 signals. Figure 6A shows the distributions of the Granger causalities between all the 110 pairs of individual signals between the two regions. In both cases (before and during the stimulus), the curves are indistinguishable and the causality factors are low. No clear relation emerges from using single time-series.

Figure 6B shows the Granger causality between region averages. Before the stimulus, the connections from left ro right and right to left have very similar distribution, with such a large error over the 40 experiments that it makes the result inconclusive. During the stimulus, the connection from right to left vanishes, while the connection from left to right significantly decreases.

In constrast, using the complex Granger causality makes for a clearer picture, as seen in figure 6C. Before stimulus, there is an almost unidirectional flux of activity from left to right. This is still the case – if less pronounced – during the stimulus. This clear assymetry between the left and right hemispheres during face recognition has been reported in the litterature not only for sheep [18, 19] and ungulates but primates as well [20].

2.5 A metabolic network

Next we turn our attention to metabolic networks. Metabolic networks consist of reactions transforming molecules of one type to molecules of another type. In modeling terms, the

concentrations of the molecules and their rates of change are of special interest. Enzymes catalyze biochemical reactions and can immensely increase the rate of a reaction. We will illustrate the effect of enzymes in this section.

The following picture is the general scheme of inhibition in Michaelis-Menten kinetics ([21, 22]). Reactions 1 and 2 belong to the standard scheme of Michaelis-Menten kinetics, *i.e.*, a reversible formation of an enzyme-substrate complex ES from the free enzyme E and the substrate S and an irreversible release of the product P from the enzyme E. Competitive inhibition is given if in addition reaction 3 (and not reaction 4,5 and 6) occurs. Uncompetitive inhibitions involves reactions 1, 2 and 4, and noncompetitive inhibition comprises reaction 1,2,3,4 and 5. Appearance of reaction 6 indicates partial inhibition.

$$\begin{array}{ccccc} E+S & \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} & ES & \xrightarrow{k_2} & E+P \\ + & & + & & \\ I & & I & & \\ {\scriptstyle k_{-3}}\Big\uparrow\Big\downarrow{\scriptstyle k_3} & & {\scriptstyle k_{-4}}\Big\uparrow\Big\downarrow{\scriptstyle k_4} & & \\ EI+S & \underset{k_5}{\overset{k_{-5}}{\rightleftarrows}} & ESI & \xrightarrow{k_6} & E+P+I \end{array}$$

In order to analyze the influence of enzyme on the reaction rate, we select a sub-reaction (reaction 2) from the above Michaelis-Menten kinetics. For simplicity, we denote the reaction as the following representation:

$$F \underset{r_{-1}}{\overset{r_1}{\rightleftarrows}} ES \xrightarrow{k_2} E+P$$

Where F represents the total substances which have influence on ES, r_1 and r_{-1} are reaction rates. Suppose there exists another enzyme E^* which only has influence on the reaction rate k_2, i.e., $E^* = ak_2$, where a is a constant. In order to measure the influence of E^*, we analyze the influence from E^* to ES, E, P, $ES+P$ and $ES-P$. The respective system of ODEs for the dynamics of this reaction reads as follows:

$$\begin{cases} \dfrac{dES}{dt} &= r_1F - r_{-1}ES - k_2ES \\ \dfrac{dE}{dt} &= k_2ES \\ \dfrac{dP}{dt} &= k_2ES \\ \dfrac{d(ES+P)}{dt} &= r_1F - r_{-1}ES \\ \dfrac{d(ES-P)}{dt} &= r_1F - r_{-1}ES - 2k_2ES \end{cases} \tag{1}$$

Furthermore, we can rewrite the above ODEs in the regression formation which have the following representations:

$$\begin{cases} ES(t) &= ES(t-1) + (r_1(t-1)F(t-1) - r_{-1}(t-1)ES(t-1) - aES(t-1)E^*(t-1))dt \\ E(t) &= E(t-1) + aES(t-1)E^*(t-1)dt \\ P(t) &= P(t-1) + aES(t-1)E^*(t-1)dt \\ (ES+P)(t) &= (ES+P)(t-1) + (r_1(t-1)F(t-1) - r_{-1}(t-1)ES(t-1))dt \\ (ES-P)(t) &= (ES-P)(t-1) + (r_1(t-1)F(t-1) - r_{-1}(t-1)ES(t-1) \\ & \quad -2aES(t-1)E^*(t-1))dt \end{cases} \tag{2}$$

From the above equations, we can see that the coefficients before E^* in the last equation is the biggest. If the enzyme E^* has a direct influence on the reaction rate r_1, then the causality from E^* to $ES - P$ should be the biggest, vice versa. We can further validate this conclusion in the following two simulation examples.

Example 4 Consider the following dynamic system:

$$\begin{cases} \frac{dS}{dt} &= -r_1 S + r_2 P_1 \\ \frac{dP_1}{dt} &= r_1 S - r_2 P_1 - r_3 P_1 + r_4 P_2 \\ \frac{dP_2}{dt} &= r_3 P_1 - r_4 P_2 \\ r_1 &= \frac{1}{1+\exp(-E)} \end{cases} \tag{3}$$

where E (enzyme) is the summation of a constant k and noise ω , $r_i, i = 1, 2, 3, 4$ are reaction rates, and r_2, r_3 and r_4 are constants, ω is a normal variable with a variance 0.05^2. From the above equations, it is clear that the structure of such a system is:

$$S \underset{r_2}{\overset{r_1}{\rightleftharpoons}} P_1 \underset{r_4}{\overset{r_3}{\rightleftharpoons}} P_2$$

Fig. 7 (**A**) demonstrates the metabolic process. There are feedbacks between S and P_1 and P_1 and P_2. Enzyme E has a direct influence on the rate r_1. We can calculate the interactions between three substances S, P_1 and P_2, Fig. 7 (**B**) is the value of partial Granger causality and their confidence interval after 1000 replications with $r_2 = 0.85, r_3 = 0.65, r_4 = 1.32, k = 0.82$. It is consistent with the true relationship. In order to analyse the influence of enzyme E, we calculate partial Granger causality from E to S, P_1, P_2, r_1, $S + P_1$ and $S - P_1$, the results are shown in Fig. 7 (**C**). It is easy to see that enzyme E has the biggest influence on $S - P_1$, but has no influence on $S + P_1$. It confirms our conclusion that if enzyme E has direct influence on the reaction rate from S to P_1, the causality from E to $S - P_1$ should be the biggest.

In Fig. 7 (**D**), we show all interactions between E and the combination of $\{S, P_1, P_2\}$. There are significant interactions from E to the group of $\{S, P_1\}$, $\{S, P_2\}$ and $\{S, P_1, P_2\}$, the results are consistent with eq.(3).

In Example 4, enzyme E has direct influence on the reaction rate from S to P_1, then E has the biggest influence on $S - P_1$. In order to comprehensively validate our approach, let us look at another example.

Example 5 Consider the following dynamic system:

$$\begin{cases} \frac{dS}{dt} &= -r_1 S + r_2 P_1 + r_5 E \\ \frac{dP_1}{dt} &= r_1 S - r_2 P_1 - r_3 P_1 + r_4 P_2 \\ \frac{dP_2}{dt} &= r_3 P_1 - r_4 P_2 \\ \frac{dE}{dt} &= -r_5 E \end{cases} \tag{4}$$

Where $r_i, i = 1, \cdots, 5$ are constants. The structure of such system is:

$$E \xrightarrow{r_5} S \underset{r_2}{\overset{r_1}{\rightleftharpoons}} P_1 \underset{r_4}{\overset{r_3}{\rightleftharpoons}} P_2$$

Fig. 7 (**E**) demonstrate the metabolic process. In this example enzyme E has direct influence on S rather than on the rate r_1. We can calculate the interactions between three substances S, P_1 and P_2, Fig. 7 (**F**) is the value of partial Granger causality and their confidence interval after 1000 replications with $r_1 = 0.8, r_2 = 1.95, r_3 = 1.85, r_4 = 1.32, r_5 = 1.8$. It is also consistent with the true relationship. Fig. 7 (**G**) is partial Granger causality from E to three substances S, P_1, P_2 and complex of $\{S, P_1, P_2\}$. From the results we can see that E has causality on $S, S+P_1, S-P_1$ and complex of S, but E has the biggest influence on S rather than $S - P_1$ (see example above). The reason is that in this example E doesn't exist a direct influence on the reaction rate r_1. It also demonstrates that if E has direct influence on the reaction rate, the causality from E to the difference of the two reactants should be the biggest.

2.6 Impact of correlation on Granger causality

Actually, the complex interaction could, more significantly, also arise from the correlation between individual units. Let us consider a model where $y_i, i = 1, 2, \cdots, N$ are identical random processes. The Granger causality from $(y_i(t), i = 1, \cdots, N)$ to $y(t) := a\sum_{i=1}^{N} y_i(t) + \epsilon_t$ is $\log(1 + a^2N(1 + \rho(N - 1)))$ where ρ is the correlation coefficient between y_i and ϵ_t is normaly distributed. Fig. 8 illustrates how the complex interaction depends on the correlation. The individual unit has interactions Fig. 8 **B**), middle panel), but as a group, it could have (Fig. 8(**B**), left panel) or have no interactions (Fig. 8 **B**)right panel). Not surprisingly, collaborative activity enhances the interaction (blue arrow in Fig. 8 **A**), but antagonistic activity reduces or even suppresses the interaction.

3 Methods

3.1 Pairwise Complex Granger causality

3.1.1 Time Domain Formulation

Consider two multiple stationary time series $\overrightarrow{X}_t$ and $\overrightarrow{Y}_t$ with k and l dimensions respectively. Individually, under fairly general conditions, each time series has the following vector autoregressive representation

$$\begin{cases} \overrightarrow{X}_t &= \sum_{i=1}^{\infty} A_{1i}\overrightarrow{X}_{t-i} + \overrightarrow{\epsilon_{1t}} \\ \overrightarrow{Y}_t &= \sum_{i=1}^{\infty} B_{1i}\overrightarrow{Y}_{t-i} + \overrightarrow{\epsilon_{2t}} \end{cases} \tag{5}$$

where $\overrightarrow{\epsilon_{it}}, i = 1, 2$ are normally distributed random vectors with k and l dimensions. Their contemporaneous covariance matrix are Γ_{xx} and Γ_{yy} with trace being denoted by T_x and T_y respectively. The value of T_x is non-negative and equals to the summation of all eigenvalues of Γ_{xx}, which measures the accuracy of the autoregressive prediction of $\overrightarrow{X}$ based on its previous values, whereas the value of T_y represents the accuracy of predicting present value of $\overrightarrow{Y}$ based on previous values of $\overrightarrow{Y}$.

Jointly, they are represented as

$$\begin{cases} \overrightarrow{X}_t &= \sum_{i=1}^{\infty} A_{2i}\overrightarrow{X}_{t-i} + \sum_{i=1}^{\infty} B_{2i}\overrightarrow{Y}_{t-i} + \overrightarrow{\epsilon_{3t}} \\ \overrightarrow{Y}_t &= \sum_{i=1}^{\infty} C_{2i}\overrightarrow{X}_{t-i} + \sum_{i=1}^{\infty} D_{2i}\overrightarrow{Y}_{t-i} + \overrightarrow{\epsilon_{4t}} \end{cases} \tag{6}$$

where the noise terms are uncorrelated over time and their contemporaneous covariance matrix is

$$\Sigma = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix} \tag{7}$$

The submatrix are defined as $\Sigma_{xx} = \text{var}(\overrightarrow{\epsilon_{3t}}), \Sigma_{xy} = \text{cov}(\overrightarrow{\epsilon_{3t}}, \overrightarrow{\epsilon_{4t}}), \Sigma_{yx} = \text{cov}(\overrightarrow{\epsilon_{4t}}, \overrightarrow{\epsilon_{3t}}), \Sigma_{yy} = \text{var}(\overrightarrow{\epsilon_{4t}})$. If $\overrightarrow{X_t}$ and $\overrightarrow{Y_t}$ are independent, the coefficient matrix B_{2i} and C_{2i} are zero, $\Gamma_{xx} = \Sigma_{xx}, \Gamma_{yy} = \Sigma_{yy}$, $\Sigma_{xy} = \Sigma_{yx}^{'} = \mathbf{0}$. The trace of Σ_{xx} and Σ_{yy} is denoted by T_{xy} and T_{yx} respectively. Consider eq. (6), the value of T_{xy} represents the accuracy of predicting present value of $\overrightarrow{X}$ based on previous values of both $\overrightarrow{X}$ and $\overrightarrow{Y}$. According to the causality definition of Granger, if the prediction of one time series is improved by incorporating past information of the second time series, then the second time series causes the first process. We extend them to multiple dimensional cases. If the trace of prediction error for the first multiple time series is reduced by the inclusion of past histories of the second multiple time series, then a causal relation from the second multiple time series to the first multiple time series exists. We quantify this causal influence by

$$F_{\overrightarrow{Y}\rightarrow\overrightarrow{X}} = \ln(T_x/T_{xy}) \tag{8}$$

It is clear that $F_{\overrightarrow{Y}\rightarrow\overrightarrow{X}} = 0$ when there is no causal influence from $\overrightarrow{X}$ to $\overrightarrow{Y}$ otherwise $F_{\overrightarrow{X}\rightarrow\overrightarrow{Y}} > 0$.

Note that in constrast with our previous extension of Granger causality ([23]), the complex Granger causality is now formulated in terms of the trace – and not the determinant – of matrices, for numerical stability and more theoretical considerations (see discussion below).

3.1.2 Frequency Domain Formulation

Time series contain oscillatory aspects in specific frequency bands. It is thus desirable to to have a spectral representation of causal influence. We then consider the frequency domain formulation of complex Granger causality. Rewrite eqs. (6) in terms of the lag operator

$$\begin{pmatrix} A_2(L) & B_2(L) \\ C_2(L) & D_2(L) \end{pmatrix} \begin{pmatrix} \overrightarrow{X}_t \\ \overrightarrow{Y}_t \end{pmatrix} = \begin{pmatrix} \overrightarrow{\epsilon_{3t}} \\ \overrightarrow{\epsilon_{4t}} \end{pmatrix} \tag{9}$$

where $A_2(0) = I_k, B_2(0) = \mathbf{0}, C_2(0) = \mathbf{0}, d_2(0) = I_l$. Fourier transforming both sides of eqs.(9) leads to

$$\begin{pmatrix} A_2(\omega) & B_2(\omega) \\ C_2(\omega) & D_2(\omega) \end{pmatrix} \begin{pmatrix} X(\omega) \\ Y(\omega) \end{pmatrix} = \begin{pmatrix} E_x(\omega) \\ E_y(\omega) \end{pmatrix} \tag{10}$$

where the components of the coefficient matrix are

$$A_2(\omega) = I_k - \sum_{i=1}^{\infty} A_{2i} e^{-i\omega j}, \;\; B_2(\omega) = -\sum_{i=1}^{\infty} B_{2i} e^{-i\omega j},$$

$$C_2(\omega) = -\sum_{i=1}^{\infty} C_{2i} e^{-i\omega j}, \;\; D_2(\omega) = I_l - \sum_{i=1}^{\infty} D_{2i} e^{-i\omega j},$$

Recasting eq.(10) into the transfer function format we obtain

$$\begin{pmatrix} X(\omega) \\ Y(\omega) \end{pmatrix} = \begin{pmatrix} H_{xx}(\omega) & H_{xy}(\omega) \\ H_{yx}(\omega) & H_{yy}(\omega) \end{pmatrix} \begin{pmatrix} E_x(\omega) \\ E_y(\omega) \end{pmatrix} \tag{11}$$

the components of $\mathbf{H}(\omega)$ are

$$H_{yy}(\omega) = (D_2(\omega) - C_2(\omega)A_2(\omega)^{-1}B_2(\omega))^{-1}$$

$$H_{xy}(\omega) = -A_2(\omega)^{-1}B_2(\omega)H_{yy}(\omega)$$

$$H_{yx}(\omega) = -H_{yy}(\omega)C_2(\omega)A_2(\omega)^{-1}$$

$$H_{xx}(\omega) = A_2(\omega)^{-1} - H_{xy}(\omega)C_2(\omega)A_2(\omega)^{-1}$$

After proper ensemble averaging we have the spectral matrix

$$S(\omega) = \mathbf{H}(\omega)\Sigma\mathbf{H}^*(\omega)$$

where * denotes the complex conjugate and matrix transpose, and Σ is defined in eq. (7).

To obtain the frequency decomposition of the time domain causality defined in the previous section, we look at the auto-spectrum of $\overrightarrow{X_t}$

$$S_{xx}(\omega) = H_{xx}(\omega)\Sigma_{xx}H^*_{xx}(\omega)+H_{xx}(\omega)\Sigma_{xy}H^*_{xy}(\omega)+H_{xy}(\omega)\Sigma_{yx}H^*_{xx}(\omega)+H_{xy}(\omega)\Sigma_{yy}H^*_{xy}(\omega) \tag{12}$$

Note that the value of diagonal of $S_{xx}(\omega)$ is real numbers, the trace of both sides can be represented as

$$\begin{aligned} \text{tr}(S_{xx}(\omega)) \;\; = \;\; & \text{tr}(H_{xx}(\omega)\Sigma_{xx}H^*_{xx}(\omega)) + \text{tr}(H_{xx}(\omega)\Sigma_{xy}H^*_{xy}(\omega) + H_{xy}(\omega)\Sigma_{yx}H^*_{xx}(\omega)) \\ & +\text{tr}(H_{xy}(\omega)\Sigma_{yy}H^*_{xy}(\omega)) \end{aligned} \tag{13}$$

We first consider a simple case $\Sigma_{xy} = \mathbf{0}$. The second term of the right side of eq.(13) is zero. We have

$$\text{tr}(S_{xx}(\omega)) = \text{tr}(H_{xx}(\omega)\Sigma_{xx}H^*_{xx}(\omega)) + \text{tr}(H_{xy}(\omega)\Sigma_{yy}H^*_{xy}(\omega)) \tag{14}$$

which implies that the spectrum of $\overrightarrow{X_t}$ has two terms, the first term, viewed as the intrinsic part, involves only the noise term that drives the $\overrightarrow{X_t}$ time series. The second term, viewed as the causal part, involves only the noise term that drives $\overrightarrow{Y_t}$.

When $\Sigma_{xy} \neq \mathbf{0}$, we can normalize eq. (10) by multiplying the following matrix

$$P = \begin{pmatrix} I_k & \mathbf{0} \\ -\Sigma_{yx}\Sigma_{xx}^{-1} & I_l \end{pmatrix} \tag{15}$$

to both sides of eq.(10). The results is

$$\begin{pmatrix} A_2(\omega) & B_2(\omega) \\ C_3(\omega) & D_3(\omega) \end{pmatrix} \begin{pmatrix} X(\omega) \\ Y(\omega) \end{pmatrix} = \begin{pmatrix} E_x(\omega) \\ \tilde{E}_y(\omega) \end{pmatrix} \tag{16}$$

where $C_3(\omega) = C_2(\omega) - \Sigma_{yx}\Sigma_{xx}^{-1}A_2(\omega), D_3(\omega) = D_2(\omega) - \Sigma_{yx}\Sigma_{xx}^{-1}B_2(\omega), \tilde{E}_y(\omega) = E_y(\omega) - \Sigma_{yx}\Sigma_{xx}^{-1}E_x(\omega)$. From the construction it is easy to see that E_x and $\tilde{E}_y$ are uncorrelated. The variance of the noise term for the normalized $\overrightarrow{Y_t}$ equation is $\Sigma_{yy} - \Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy}$. The new transfer function $\tilde{\mathbf{H}}(\omega)$ for eq.(16) is the inverse of the new coefficient matrix

$$\tilde{\mathbf{H}}(\omega) = \begin{pmatrix} \tilde{H}_{xx}(\omega) & \tilde{H}_{xy}(\omega) \\ \tilde{H}_{yx}(\omega) & \tilde{H}_{yy}(\omega) \end{pmatrix} \tag{17}$$

where

$$\tilde{H}_{xx}(\omega) = H_{xx}(\omega) + H_{xy}(\omega)\Sigma_{yx}\Sigma_{xx}^{-1}, \quad \tilde{H}_{xy}(\omega) = H_{xy}(\omega)$$
$$\tilde{H}_{yx}(\omega) = H_{yx}(\omega) + H_{yy}(\omega)\Sigma_{yx}\Sigma_{xx}^{-1}, \quad \tilde{H}_{yy}(\omega) = H_{yy}(\omega)$$

Note that E_x and $\tilde{E}_y$ are uncorrelated, following the same steps of eq.(14), the spectrum of $\overrightarrow{X_t}$ is found to be

$$S_{xx}(\omega) = \tilde{H}_{xx}(\omega)\Sigma_{xx}\tilde{H}_{xx}^*(\omega) + \tilde{H}_{xy}(\omega)(\Sigma_{yy} - \Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy})\tilde{H}_{xy}^*(\omega) \tag{18}$$

The trace of both sides can be represented as

$$\text{tr}(S_{xx}(\omega)) = \text{tr}(\tilde{H}_{xx}(\omega)\Sigma_{xx}\tilde{H}_{xx}^*(\omega)) + \text{tr}(\tilde{H}_{xy}(\omega)(\Sigma_{yy} - \Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy})\tilde{H}_{xy}^*(\omega)) \tag{19}$$

Here he first term is interpreted as the intrinsic power and the second term as the causal power of $\overrightarrow{X_t}$ due to $\overrightarrow{Y_t}$. We define the causal influence from $\overrightarrow{Y_t}$ to $\overrightarrow{X_t}$ at frequency ω as

$$f_{\overrightarrow{Y}\to\overrightarrow{X}}(\omega) = \ln \frac{\text{tr}(S_{xx}(\omega))}{\text{tr}(\tilde{H}_{xx}(\omega)\Sigma_{xx}\tilde{H}_{xx}^*(\omega))} \tag{20}$$

3.2 Partial Complex Granger causality

3.2.1 Time Domain Formulation

Consider three multiple stationary time series $\overrightarrow{X_t}, \overrightarrow{Y_t}$ and $\overrightarrow{Z_t}$ with k, l and m dimensions respectively. We first consider the relationship from $\overrightarrow{Y_t}$ to $\overrightarrow{X_t}$ on condition of $\overrightarrow{Z_t}$. The joint autoregressive representation for $\overrightarrow{X_t}$ and $\overrightarrow{Z_t}$ can be written as

$$\begin{cases} \overrightarrow{X_t} &= \sum_{i=1}^{\infty} a_{1i}\overrightarrow{X}_{t-i} + \sum_{i=1}^{\infty} c_{1i}\overrightarrow{Z}_{t-i} + \overrightarrow{\epsilon_{1t}} \\ \overrightarrow{Z_t} &= \sum_{i=1}^{\infty} b_{1i}\overrightarrow{Z}_{t-i} + \sum_{i=1}^{\infty} d_{1i}\overrightarrow{X}_{t-i} + \overrightarrow{\epsilon_{2t}} \end{cases} \tag{21}$$

The noise covariance matrix for the system can be represented as

$$\Gamma = \begin{pmatrix} \text{var}(\overrightarrow{\epsilon_{1t}}) & \text{cov}(\overrightarrow{\epsilon_{1t}}, \overrightarrow{\epsilon_{2t}}) \\ \text{cov}(\overrightarrow{\epsilon_{2t}}, \overrightarrow{\epsilon_{1t}}) & \text{var}(\overrightarrow{\epsilon_{2t}}) \end{pmatrix} = \begin{pmatrix} \Gamma_{xx} & \Gamma_{xz} \\ \Gamma_{zx} & \Gamma_{zz} \end{pmatrix}$$

where var and cov represent variance and covariance respectively. Extending this representation, the vector autoregressive representation for a system involving three time series $\overrightarrow{X_t}$, $\overrightarrow{Y_t}$ and $\overrightarrow{Z_t}$ can be written in the following way.

$$\begin{cases} \overrightarrow{X_t} &= \sum_{i=1}^{\infty} a_{2i}\overrightarrow{X}_{t-i} + \sum_{i=1}^{\infty} b_{2i}\overrightarrow{Y}_{t-i} + \sum_{i=1}^{\infty} c_{2i}\overrightarrow{Z}_{t-i} + \overrightarrow{\epsilon_{3t}} \\ \overrightarrow{Y_t} &= \sum_{i=1}^{\infty} d_{2i}\overrightarrow{X}_{t-i} + \sum_{i=1}^{\infty} e_{2i}\overrightarrow{Y}_{t-i} + \sum_{i=1}^{\infty} f_{2i}\overrightarrow{Z}_{t-i} + \overrightarrow{\epsilon_{4t}} \\ \overrightarrow{Z_t} &= \sum_{i=1}^{\infty} g_{2i}\overrightarrow{X}_{t-i} + \sum_{i=1}^{\infty} h_{2i}\overrightarrow{Y}_{t-i} + \sum_{i=1}^{\infty} k_{2i}\overrightarrow{Z}_{t-i} + \overrightarrow{\epsilon_{5t}} \end{cases} \tag{22}$$

The noise covariance matrix for the above system can be represented as

$$\Sigma = \begin{pmatrix} \text{var}(\overrightarrow{\epsilon_{3t}}) & \text{cov}(\overrightarrow{\epsilon_{3t}}, \overrightarrow{\epsilon_{4t}}) & \text{cov}(\overrightarrow{\epsilon_{3t}}, \overrightarrow{\epsilon_{5t}}) \\ \text{cov}(\overrightarrow{\epsilon_{4t}}, \overrightarrow{\epsilon_{3t}}) & \text{var}(\overrightarrow{\epsilon_{4t}}) & \text{cov}(\overrightarrow{\epsilon_{4t}}, \overrightarrow{\epsilon_{5t}}) \\ \text{cov}(\overrightarrow{\epsilon_{5t}}, \overrightarrow{\epsilon_{3t}}) & \text{cov}(\overrightarrow{\epsilon_{5t}}, \overrightarrow{\epsilon_{4t}}) & \text{var}(\overrightarrow{\epsilon_{5t}}) \end{pmatrix} = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} & \Sigma_{xz} \\ \Sigma_{yx} & \Sigma_{yy} & \Sigma_{yz} \\ \Sigma_{zx} & \Sigma_{zy} & \Sigma_{zz} \end{pmatrix}$$

where $\overrightarrow{\epsilon_{it}}, i = 1, \cdots, 5$ are the prediction error, which are uncorrelated over time. The conditional variance $\Gamma_{xx} - \Gamma_{xz}\Gamma_{zz}^{-1}\Gamma_{zx}$ measures the accuracy of the autoregressive prediction of $\overrightarrow{X}$ based on its previous values conditioned on $\overrightarrow{Z}$ whereas the conditional variance $\Sigma_{xx} - \Sigma_{xz}\Sigma_{zz}^{-1}\Sigma_{zx}$ measures the accuracy of the autoregressive prediction of $\overrightarrow{X}$ based on its previous values of both $\overrightarrow{X}$ and $\overrightarrow{Y}$ conditioned on $\overrightarrow{Z}$. The trace of matrix $\Gamma_{xx} - \Gamma_{xz}\Gamma_{zz}^{-1}\Gamma_{zx}$ and matrix $\Sigma_{xx} - \Sigma_{xz}\Sigma_{zz}^{-1}\Sigma_{zx}$ are denoted by $T_{x|z}$ and $T_{xy|z}$ respectively. We define the partial Granger causality from vector $\overrightarrow{Y}$ to vector $\overrightarrow{X}$ conditioned on vector $\overrightarrow{Z}$ to be

$$F_{\overrightarrow{Y} \to \overrightarrow{X}|\overrightarrow{Z}} = \ln\left(\frac{T_{x|z}}{T_{xy|z}}\right) \tag{23}$$

3.2.2 Frequency Domain Formulation

To derive the spectral decomposition of the time domain partial Granger causality, we first multiply the matrix

$$P_1 = \begin{pmatrix} I_k & -\Gamma_{xz}\Gamma_{zz}^{-1} \\ \mathbf{0} & I_m \end{pmatrix} \tag{24}$$

to both sides of eq. (21). The normalized equations are represented as:

$$\begin{pmatrix} D_{11}(L) & D_{12}(L) \\ D_{21}(L) & D_{22}(L) \end{pmatrix} \begin{pmatrix} \overrightarrow{X_t} \\ \overrightarrow{Z_t} \end{pmatrix} = \begin{pmatrix} \tilde{X}_t \\ \tilde{Z}_t \end{pmatrix} \tag{25}$$

with $D_{11}(0) = I_k, D_{22}(0) = I_m, D_{21}(0) = \mathbf{0}, \text{cov}(\tilde{X}_t, \tilde{Z}_t) = 0$, we note that $\text{var}(\tilde{X}_t) = \Gamma_{xx} - \Gamma_{xz}\Gamma_{zz}^{-1}\Gamma_{zx}, \text{var}(\tilde{Z}_t) = \Gamma_{zz}$. For eq. (22), we also multiply the matrix

$$P = P_3 \cdot P_2 \tag{26}$$

where

$$P_2 = \begin{pmatrix} I_k & \mathbf{0} & -\Sigma_{xz}\Sigma_{zz}^{-1} \\ \mathbf{0} & I_l & -\Sigma_{yz}\Sigma_{zz}^{-1} \\ \mathbf{0} & \mathbf{0} & I_m \end{pmatrix} \tag{27}$$

and

$$P_3 = \begin{pmatrix} I_k & \mathbf{0} & \mathbf{0} \\ -(\Sigma_{yx} - \Sigma_{yz}\Sigma_{zz}^{-1}\Sigma_{zx})(\Sigma_{xx} - \Sigma_{xz}\Sigma_{zz}^{-1}\Sigma_{zx})^{-1} & I_l & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & I_m \end{pmatrix} \tag{28}$$

to both sides of eq.(22). The normalized equation of eq. (22) becomes

$$\begin{pmatrix} B_{11}(L) & B_{12}(L) & B_{13}(L) \\ B_{21}(L) & B_{22}(L) & B_{23}(L) \\ B_{31}(L) & B_{32}(L) & B_{33}(L) \end{pmatrix} \begin{pmatrix} \overrightarrow{X_t} \\ \overrightarrow{Y_t} \\ \overrightarrow{Z_t} \end{pmatrix} = \begin{pmatrix} \epsilon_{xt} \\ \epsilon_{yt} \\ \epsilon_{zt} \end{pmatrix} \tag{29}$$

where $\epsilon_{xt}, \epsilon_{yt}, \epsilon_{zt}$ are independent, and their variances being $\hat{\Sigma}_{xx}, \hat{\Sigma}_{yy}$ and $\hat{\Sigma}_{zz}$ with

$$\begin{cases} \hat{\Sigma}_{zz} &= \Sigma_{zz} \\ \hat{\Sigma}_{xx} &= \Sigma_{xx} - \Sigma_{xz}\Sigma_{zz}^{-1}\Sigma_{zx} \\ \hat{\Sigma}_{yy} &= \Sigma_{yy} - \Sigma_{yz}\Sigma_{zz}^{-1}\Sigma_{zy} - \dfrac{(\Sigma_{yx} - \Sigma_{yz}\Sigma_{zz}^{-1}\Sigma_{zx})(\Sigma_{xy} - \Sigma_{xz}\Sigma_{zz}^{-1}\Sigma_{zy})}{(\Sigma_{yy} - \Sigma_{yz}\Sigma_{zz}^{-1}\Sigma_{zy})} \end{cases}$$

After Fourier transforming eq. (25) and eq. (29), we can rewrite these two equations in the following expression:

$$\begin{pmatrix} X(\omega) \\ Z(\omega) \end{pmatrix} = \begin{pmatrix} G_{xx}(\omega) & G_{xz}(\omega) \\ G_{zx}(\omega) & G_{zz}(\omega) \end{pmatrix} \begin{pmatrix} \tilde{X}(\omega) \\ \tilde{Z}(\omega) \end{pmatrix} \tag{30}$$

and

$$\begin{pmatrix} X(\omega) \\ Y(\omega) \\ Z(\omega) \end{pmatrix} = \begin{pmatrix} H_{xx}(\omega) & H_{xy}(\omega) & H_{xz}(\omega) \\ H_{yx}(\omega) & H_{yy}(\omega) & H_{yz}(\omega) \\ H_{zx}(\omega) & H_{zy}(\omega) & H_{zz}(\omega) \end{pmatrix} \begin{pmatrix} E_x(\omega) \\ E_y(\omega) \\ E_z(\omega) \end{pmatrix} \tag{31}$$

Note that $X(\omega)$ and $Z(\omega)$ from eq. (30) are identical with that from eq. (31), we thus have

$$\begin{aligned} \begin{pmatrix} \tilde{X}(\omega) \\ Y(\omega) \\ \tilde{Z}(\omega) \end{pmatrix} &= \begin{pmatrix} G_{xx}(\omega) & 0 & G_{xz}(\omega) \\ 0 & 1 & 0 \\ G_{zx}(\omega) & 0 & G_{zz}(\omega) \end{pmatrix}^{-1} \begin{pmatrix} H_{xx}(\omega) & H_{xy}(\omega) & H_{xz}(\omega) \\ H_{yx}(\omega) & H_{yy}(\omega) & H_{yz}(\omega) \\ H_{zx}(\omega) & H_{zy}(\omega) & H_{zz}(\omega) \end{pmatrix} \begin{pmatrix} E_x(\omega) \\ E_y(\omega) \\ E_z(\omega) \end{pmatrix} \\ &= \begin{pmatrix} Q_{xx}(\omega) & Q_{xy}(\omega) & Q_{xz}(\omega) \\ Q_{yx}(\omega) & Q_{yy}(\omega) & Q_{yz}(\omega) \\ Q_{zx}(\omega) & Q_{zy}(\omega) & Q_{zz}(\omega) \end{pmatrix} \begin{pmatrix} E_x(\omega) \\ E_y(\omega) \\ E_z(\omega) \end{pmatrix} \end{aligned} \tag{32}$$

where $\mathbf{Q}(\omega) = \mathbf{G}^{-1}(\omega)\mathbf{H}(\omega)$. Now the power spectrum of $\tilde{X}$ is

$$S_{\tilde{x}\tilde{x}}(\omega) = Q_{xx}(\omega)\hat{\Sigma}_{xx}Q_{xx}^*(\omega) + Q_{xy}(\omega)\hat{\Sigma}_{yy}Q_{xy}^*(\omega) + Q_{xz}(\omega)\hat{\Sigma}_{zz}Q_{xz}^*(\omega) \tag{33}$$

The trace of both sides of eq. (33) can be represented as

$$\text{tr}(S_{\tilde{x}\tilde{x}}(\omega)) = \text{tr}(Q_{xx}(\omega)\hat{\Sigma}_{xx}Q_{xx}^*(\omega)) + \text{tr}(Q_{xy}(\omega)\hat{\Sigma}_{yy}Q_{xy}^*(\omega)) + \text{tr}(Q_{xz}(\omega)\hat{\Sigma}_{zz}Q_{xz}^*(\omega)) \tag{34}$$

Note that $\hat{\Sigma}_{xx} = \Sigma_{xx} - \Sigma_{xz}\Sigma_{zz}^{-1}\Sigma_{zx}$, the first term can be thought of as the intrinsic power eliminating exogenous inputs and latent variables and the remaining two terms as the combined causal influence from $\overrightarrow{Y}$ on the mediate of $\overrightarrow{Z}$. This interpretation leads immediately to the definition

$$f_{\overrightarrow{Y}\to\overrightarrow{X}|\overrightarrow{Z}}(\omega) = \ln \frac{\text{tr}(S_{\tilde{x}\tilde{x}}(\omega))}{\text{tr}(Q_{xx}(\omega)\hat{\Sigma}_{xx}Q_{xx}^{*}(\omega))} \tag{35}$$

In previous studies, we showed that by the Kolmogrov formula [11] for spectral decompositions and under some mild conditions, the Granger causality in the frequency domain and in the time domain satisfy

$$F_{\overrightarrow{Y}\to\overrightarrow{X}|\overrightarrow{Z}} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_{\overrightarrow{Y}\to\overrightarrow{X}|\overrightarrow{Z}}(\omega)d\omega \tag{36}$$

All our numerical simulations and applications on real data strongly suggest this is still the case with the present extension of the definition. However, whether it is true in general remains a conjecture at this stage.

4 Discussion

We have presented a study for the complex Granger causality. The time domain complex Granger causality and its frequency domain decomposition have been successfully applied to simulated examples and experimental data.

4.1 An improvement over Partial Granger Causality

In [23], we have introduced the notion of partial Granger causality and successfully applied it to gene and neuron data. Although partial Granger causality is formulated for any dimension (see [23] Eq. 5), there are problems when we face high dimensional data and we actually only restricted ourselves to one dimensional case. The partial Granger causality is defined as the ratio of two determinants and as we know

$$|A| = \prod \lambda_i$$

where λ_i are all eigenvalues of the matrix which in theory should be positive definite, but in practice we often have the case that some of the eigenvalues could be zero (positive semidefinite). Hence when we use the partial Granger causality defined in [23] to deal with real data, the partial Granger causality is not well defined. At the present paper, we define the partial Granger causality as the trace of their variance matrix and it is certainly more stable in numerical examples. Nevertheless, we want to stress that when we deal with one dimensional case, the Granger causality defined in terms of the determinant or trace is identical.

In the one dimensional case, the Granger causality is always non-negative. Can we generalize this property to complex interactions as we developed here? In other words, can we introduce an order in the space of all variance matrices. Actually, as in the literature (see for example, page 469 of [24]) we can define a variance matrix A is greater than B, denoting as $A \succ B$ if and only if $A - B$ is positive semidefinite. However, we can easily see that Y is a complex Granger cause of X does not imply that $\Gamma_{xx} \succ \Sigma_{xx}$.

4.2 The importance of considering complex interactions

If we want to understand biological processes in details, at least two things are required: a large amount of accurate data and suitable computational tools to exploit them. Thanks to the continuing advances of bio-technology, we are now in a situation where a wealth of data is not only routinely acquired but also easily available (*e.g.* [25, 26] for microarray experiments, [27] for neurophysiology). Moreover, this trend is accelerating, with new technologies becoming available ([28, 29]). The challenge now is to develop the tools necessary to make use of this information.

One approach to uncover the relations between elements of a system is to use the statistical properties of the measurements to infer 'functional' ([30]) connectivity. This is the case for, for example, Bayesian networks ([31]), Dynamic Bayesian networks ([32]) or Granger causality networks ([33]. Typically, a global network is inferred from the connectivity from one element to another, or from one group of elements ('parents' in Bayesian Network settings) to a single one. This approach has produced informative results ([34, 35]) and is a very active domain of research.

But there is a real need now to go one step further, beyond these types of interactions, and to be able to deal with more complex interactions to reveal the influence of an element on the connection between two others for example, or to detect group-to-group interactions. Such complex interactions are ubiquitous in biological processes: enzymes act on the production rate of metabolites, information is passed on from one layer of neurons to the next, transcription factors form complexes which influence gene activity etc. And such interactions will be missed out with current methods.

In this paper, we demonstrated that the newly defined complex Granger Causality is able to capture these kinds of connections. For example, we showed that considering the effect of transcription factors improves network inferences in the case of the yeast-cell cycle data (Fig. 5). The method was also able to detect the effect of the enzyme in a metabolic reaction (Fig. 7) and to give a clearer and more principled picture of brain area interactions than simple averaging (Fig. 6). Having defined a measure to quantify these processes is a crucial step towards deducing the complete mechanism of a biological system. The next challenge, however, is to come: how to define the correct/relevant grouping.

4.3 Future challenges for systems biology

Consider a network of N units (genes, proteins, neurons etc.). We intend to reveal all interactions in the network, this is the driving force behind the current systems biology approach ([36]). The belief is that the network interactions are the key for understanding many meaningful biology functions: from various diseases to brain function. For a network of N units, we might plausibly assume that there are N^2 pairwise interactions (including self-interactions). Furthermore, a biological network is usually sparse and the total number of interactions should be much smaller. Hence, with simultaneously recorded data at N units, we hope to be able to recover all interactions. Here we point out, however, that when using synthesized and biological data, the number of actual interactions should be of order $\exp(N)$, since all possible subsets (complexes) of $1, 2, \cdots, N$ should be taken into account. This leads to an NP hard problem and a direct approach is bound to fail to reveal

all interactions. The search space is now much bigger: we are not looking for the correct directed acyclic graph, or even graph but the correct hypergraph. Is a systems biology approach which requires to reveal all interactions including complex interactions reported here really feasible?

4.4 Acknowledgments

We would like to thank Yang Zhang for providing us with the local field potential recordings.

References

[1] Eichler M (2006) Graphical modelling of dynamic relationships in multivariate time series. In: Schelter B, Winterhalder M, Timmer J, editors, Handbook of Time Series Analysis, Wiley-VCH Verlage. pp. 335–372.

[2] Granger CWJ (1969) Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37:424–438. doi:10.2307/1912791.

[3] Granger C (1980) Testing for causality a personal viewpoint. Journal of Economic Dynamics and Control 2:329–352. doi:10.1016/0165-1889(80)90069-X.

[4] Pearl J (2000) Causality : Models, Reasoning, and Inference. Cambridge University Press.

[5] Gourévitch B, Bouquin-Jeannès R, Faucon G (2006) Linear and nonlinear causality between signals: methods, examples and neurophysiological applications. Biological Cybernetics 95:349–369. doi:10.1007/s00422-006-0098-0.

[6] Wu J, Liu X, Feng J (2008) Detecting causality between different frequencies. Journal of Neuroscience Methods 167:367–375. doi:10.1016/j.jneumeth.2007.08.022.

[7] Wiener N (1956) The theory of prediction. In: Beckenbach EF, editor, Modern mathematics for engineers, McGraw-Hill, New York, chapter 8.

[8] Chen Y, Bressler SL, Ding M (2006) Frequency decomposition of conditional granger causality and application to multivariate neural field potential data. Journal of Neuroscience Methods 150:228–237. doi:10.1016/j.jneumeth.2005.06.011.

[9] Ding M, Chen Y, Bressler SL (2006) Granger causality: Basic theory and application to neuroscience. In: Schelter B, Winterhalder M, Timmer J, editors, Handbook of Time Series Analysis: Recent Theoretical Developments and Applications, Wiley-VCH, chapter 17.

[10] Geweke J (1982) Measurement of linear dependence and feedback between multiple time series. Journal of the American Statistical Association 77:304–313.

[11] Geweke JF (1984) Measures of conditional linear dependence and feedback between time series. Journal of the American Statistical Association 79:907–915.

[12] Spellman PT, Sherlock G, Zhang MQ, Iyer VR, Anders K, et al. (1998) Comprehensive identification of cell cycle-regulated genes of the yeast saccharomyces cerevisiae by microarray hybridization. Mol Biol Cell 9:3273–3297.

[13] Wang RS, Wang Y, Zhang XS, Chen L (2007) Inferring transcriptional regulatory networks from high-throughput data. Bioinformatics doi: 10.1093/bioinformatics/btm465.

[14] Mewes HW, Frishman D, Güldener U, Mannhaupt G, Mayer K, et al. (2002) Mips: a database for genomes and protein sequences. Nucleic Acids Res 30:31–34. doi: 10.1093/nar/30.1.31.

[15] Teixeira MC, Monteiro P, Jain P, Tenreiro S, Fernandes AR, et al. (2006) The YEASTRACT database: a tool for the analysis of transcription regulatory associations in saccharomyces cerevisiae. Nucleic Acids Res 34. doi:10.1093/nar/gkj013.

[16] Royer L, Reimann M, Andreopoulos B, Schroeder M (2008) Unraveling protein networks with power graph analysis. PLoS Comput Biol 4:e1000108+. doi: 10.1371/journal.pcbi.1000108.

[17] David O, Guillemain I, Saillet S, Reyt S, Deransart C, et al. (2008) Identifying neural drivers with functional mri: An electrophysiological validation. PLoS Biology 6:e315+. doi:10.1371/journal.pbio.0060315.

[18] Tate AJ, Fischer H, Leigh AE, Kendrick KM (2006) Behavioural and neurophysiological evidence for face identity and face emotion processing in animals. Philos Trans R Soc Lond B Biol Sci 361:2155–2172. doi:10.1098/rstb.2006.1937.

[19] Peirce J, Leigh AE, Kendrick KM (2000) Configurational coding, familiarity and the right hemisphere advantage for face recognition in sheep. Neuropsychologia 38:475–483. doi:10.1016/S0028-3932(99)00088-3.

[20] Kosslyn SM, Gazzaniga MS, Galaburda AM, Rabin C (1998) Hemispheric specialization. In: Zigmond MJ, Bloom FE, Landis SC, Roberts JL, Squire LR, editors, Fundamental Neuroscience, Academic Press Inc, chapter 58. pp. 1521–1542.

[21] Nelson DL, Cox MM (2004) Lehninger Principles of Biochemistry, Fourth Edition. W. H. Freeman.

[22] Klipp E, Herwig R, Kowald A, Wierling C, Lehrach H (2005) Systems Biology in Practice: Concepts, Implementation and Application. Wiley-VCH.

[23] Guo S, Wu J, Ding M, Feng J (2008) Uncovering interactions in the frequency domain. PLoS Comput Biol 4:e1000087+. doi:10.1371/journal.pcbi.1000087.

[24] Horn RA, Johnson CR (1990) Matrix Analysis. Cambridge University Press.

[25] Parkinson H, Kapushesky M, Kolesnikov N, Rustici G, Shojatalab M, et al. (2008) Arrayexpress update–from an archive of functional genomics experiments to the atlas of gene expression. Nucl Acids Res :gkn889+doi:10.1093/nar/gkn889.

[26] Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, et al. (2007) Ncbi geo: mining tens of millions of expression profilesdatabase and tools update. Nucleic Acids Research 35:D760–D765. doi:10.1093/nar/gkl887.

[27] Fletcher M, Liang B, Smith L, Knowles A, Jackson T, et al. (2008) Neural network based pattern matching and spike detection tools and services in the CARMEN neuroinformatics project. Neural Networks 21:1076–1084. doi:10.1016/j.neunet.2008.06.009.

[28] Kahvejian A, Quackenbush J, Thompson JF (2008) What would you do if you could sequence everything? Nat Biotech 26:1125–1133. doi:10.1038/nbt1494.

[29] Shendure J (2008) The beginning of the end for microarrays? Nat Meth 5:585–587. doi:10.1038/nmeth0708-585.

[30] Friston K (2009) Causal modelling and brain connectivity in functional magnetic resonance imaging. PLoS Biology 7:e33+. doi:10.1371/journal.pbio.1000033.

[31] Needham CJ, Bradford JR, Bulpitt AJ, Westhead DR (2007) A primer on learning in bayesian networks for computational biology. PLoS Computational Biology 3:e129+. doi:10.1371/journal.pcbi.0030129.

[32] Yu J, Smith AV, Wang PP, Hartemink AJ (2004) Advances to bayesian network inference for generating causal networks from observational biological data. Bioinformatics 20:3594+. doi:10.1093/bioinformatics/bth448.

[33] Zou C, Feng J (2009) Granger causality vs dynamic bayesian network inference: a comparative study.

[34] Sachs K, Perez O, Pe'er D, Lauffenburger DA, Nolan GP (2005) Causal protein-signaling networks derived from multiparameter single-cell data. Science 308:523–529. doi:10.1126/science.1105809.

[35] Mukherjee S, Speed TP (2008) Network inference using informative priors. Proceedings of the National Academy of Sciences 105:14313–14318. doi:10.1073/pnas.0802272105.

[36] Aebersold R, Hood LE, Watts JD (2000) Equipping scientists for the new biology. Nat Biotech 18:359. doi:10.1038/74325.

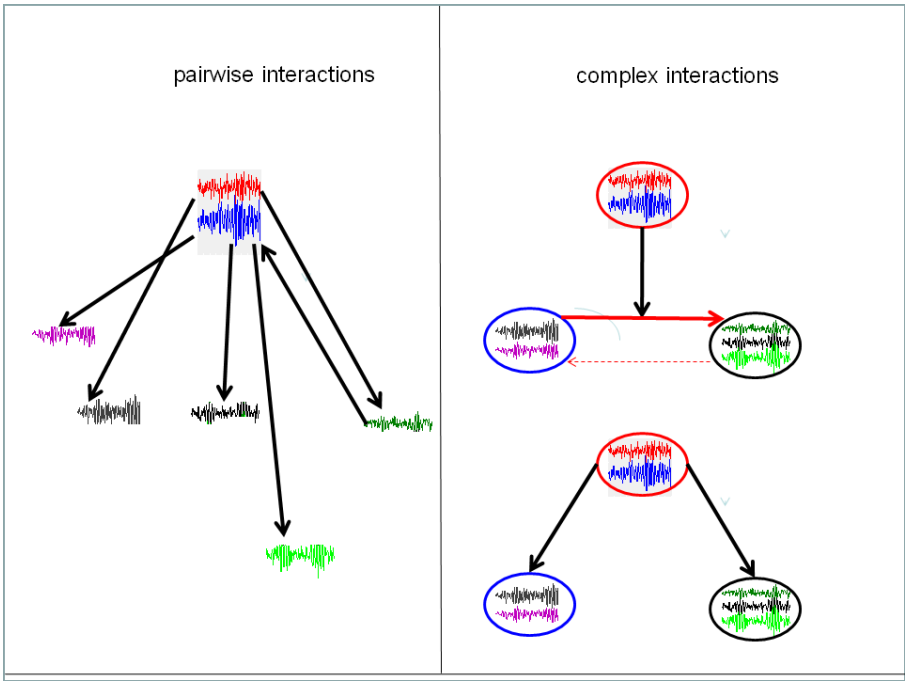

Figure 1: A schematic plot of the complex interactions. Each time trace (node) is the activity of a gene, protein, substance etc. A circle is a complex comprising of nodes. Left panel is the interactions among nodes. Right panel, the top complex can exert its influence on the rate between two complexes (top), or on the complexes themselves (bottom).

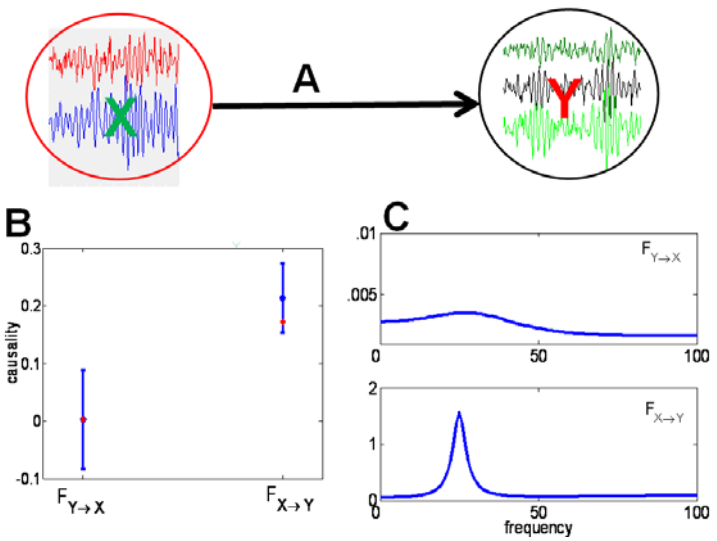

Figure 2: (**A**). Traces of the time series considered in Example 1 (inside ovals). The underlying causal relationship is represented by arrow. (**B**) Comparison between the time domain pairwise Granger causality and the frequency domain pairwise Granger causality. Blue line is the value and its confidence intervals of the pairwise complex Granger causality calculated in the time domain. Red dots are the summation (integration) of the pairwise complex Granger causality for frequencies in the range of $[-\pi, \pi]$. (**C**) are the results obtained in the frequency domain.

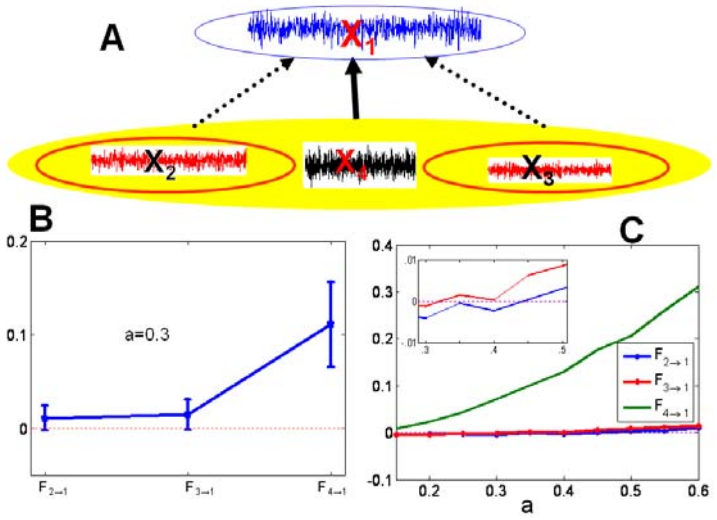

Figure 3: (**A**) The time traces of $X_i(t), i = 1, 2, 3, 4$ with $a = 0.3$. (**B**) The exact value and their confidence interval of the Granger causality in example 2 when $a = 0.3$. There are no causal relations between X_2 and X_1, and X_3 and X_1, but the causal relationship between X_4 and X_1 is significant. (**C**) The lower bound of the causal relationship as a function of a. Inset clearly shows that when a increase, the Granger causality between X_3 and X_1 and X_2 and X_1 becomes significant respectively.

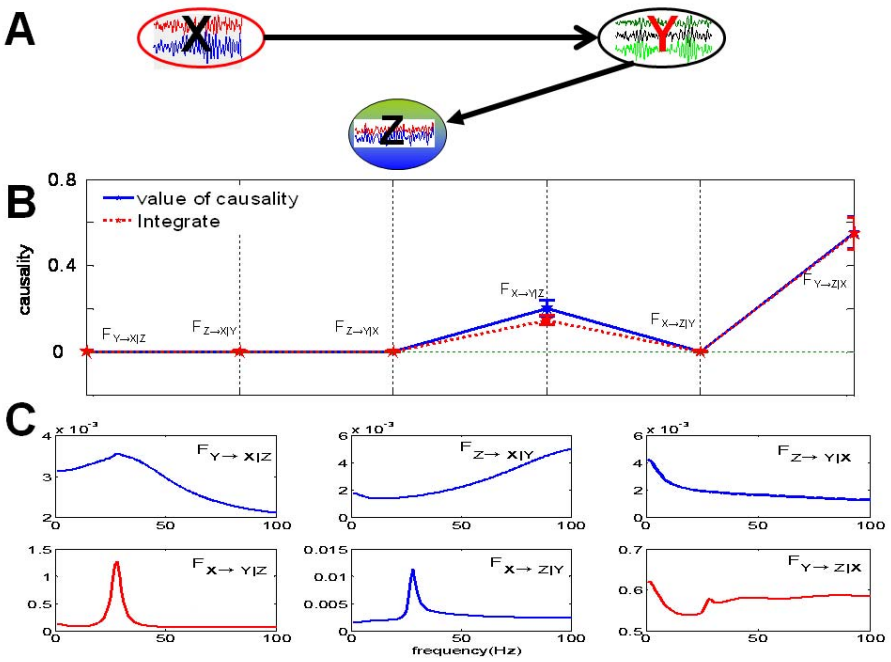

Figure 4: (**A**). Traces of the time series and the underlying causal relationships considered in example 3, both X, Y and Z are complexes (in circles). (**B**) The blue solid line is the partial complex Granger causality and its 95% confidence intervals after 1000 replications. The red dotted line is the summation (integration) of the partial complex Granger causality for frequencies in the range of $[-\pi, \pi]$. (**C**) shows the results obtained in the frequency domain.

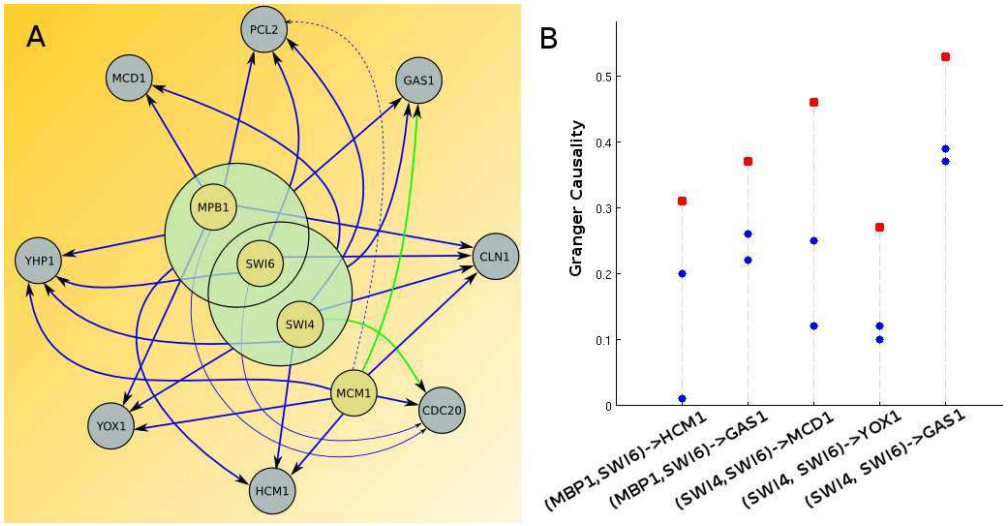

Figure 5: **A**) Inferred regulatory network of 12 genes known to participate in the yeast cell-cycle. Thick lines (blue if the interaction is documented, green if only potential according to YEASTRACT) are correct inferences. Thin lines are wrong inferences, with a dashed line representing a missed connection and a solid line representing a wrongly attributed connection. Yellow nodes denote target factors, green nodes complexes and blue nodes target genes. **B**) Improvement of the connection when complexes are considered. Blue dots represent the Granger causality from one member of the complex to the target gene, red squares represent the Granger causality from the complex to the target gene. Note that this hypergraph is not to be read as a power graph ([16]) as a connection from a complex to a target gene does not imply significant interactions from each of the subset elements to the target.

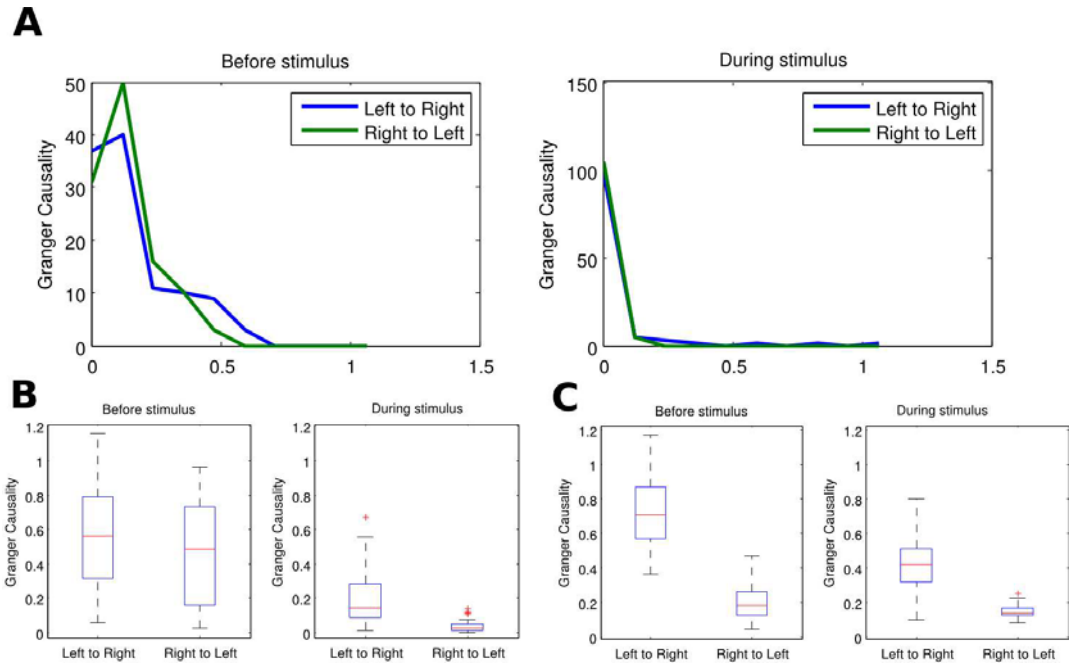

Figure 6: **A**) Distribution of Granger causality between all 110 pairs of left and right signals. **B**) Granger causality between region averages for each of the 40 experiments. **C**) Complex Granger causality between the two regions for each of the 40 experiments.

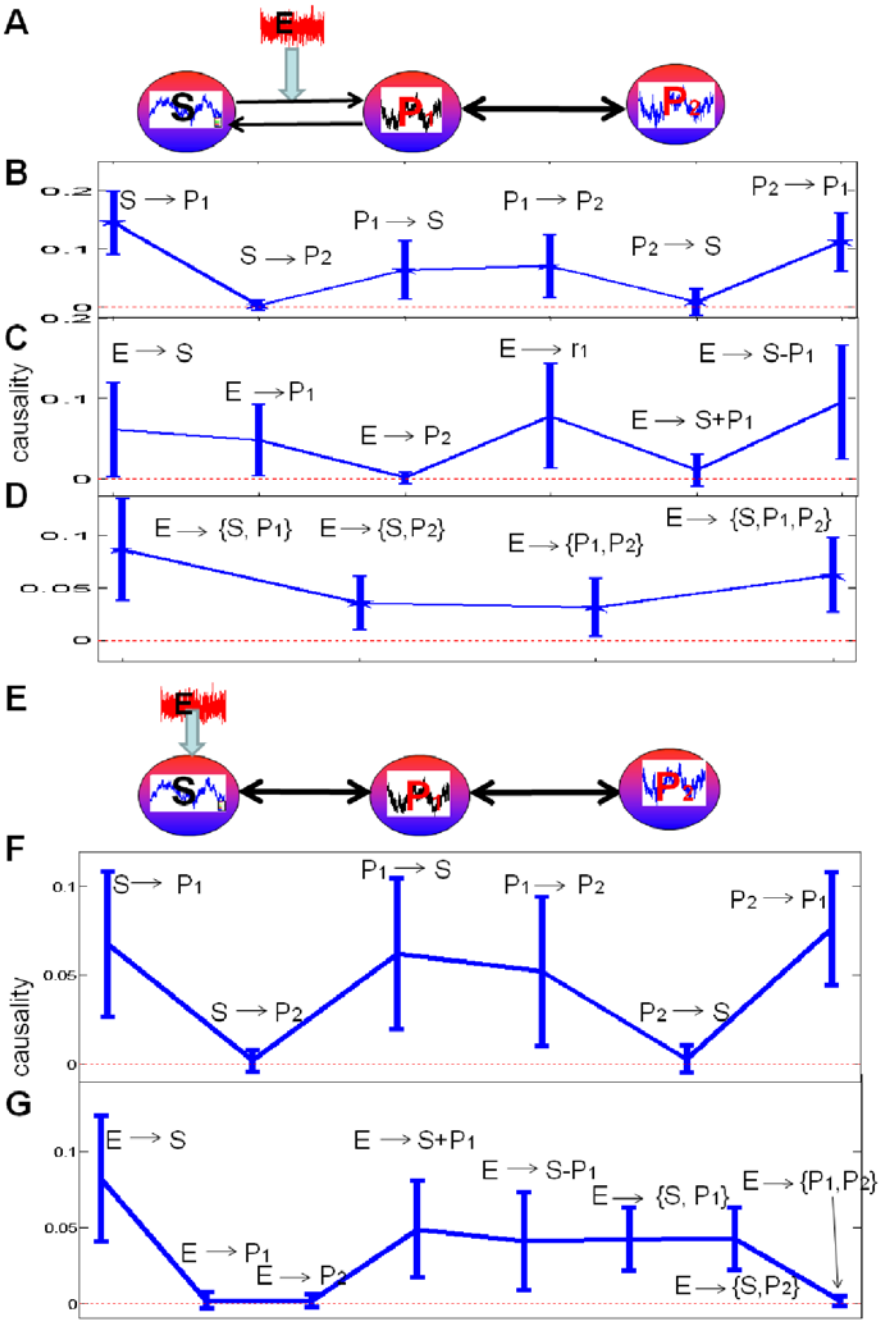

Figure 7: (**A**): trace of three reactant S, P_1 and P_2 in Example 4, the enzyme E has direct influence on the reaction rate r_1. (**B**): Partial Granger causality and its confidence interval after 1000 replications between three substance S, P_1 and P_2 in Example 4. (**C**): Partial Granger causality from E to other substances in Example 4. (**D**): Partial Granger causality from E to complex of $\{S, P_1, P_2\}$. in Example 4. (**E**): trace of three reactant S, P_1 and P_2 in Example 5, in this Example, enzyme E has direct influence on E. (**F**): Partial Granger causality and its confidence interval after 1000 replications between three substance S, P_1 and P_2 in Example 5. (**G**): Partial Granger causality from E to other substances and complex of $\{S, P_1, P_2\}$ in Example 5.

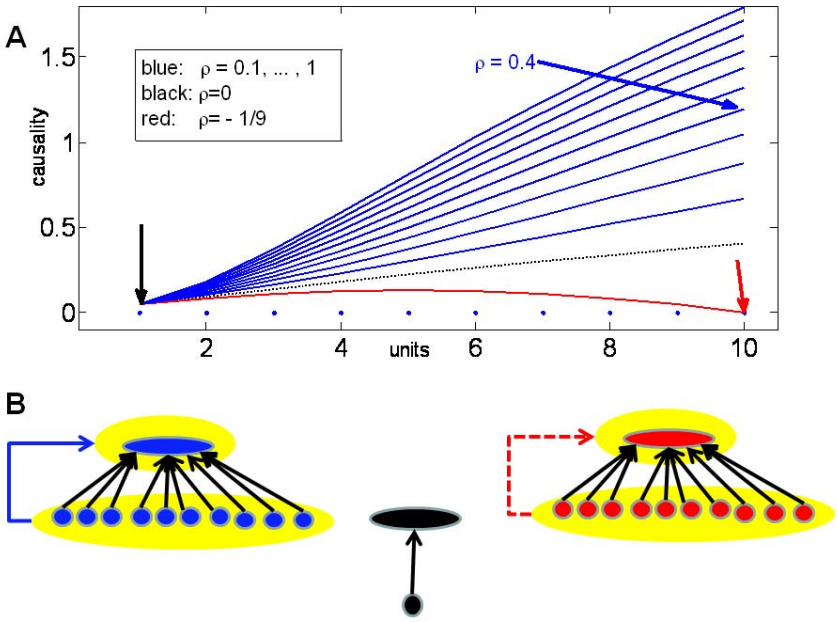

Figure 8: The role of correlation in the complex interaction. A. The Granger causality vs. units (N) for different with a=0.022. B. Three network structures (blue, black and red) correspond to the three arrows in A. Solid line indicates the existence of an interaction, but dashed line not.