New Journal of Physics

Characterizing time series: when Granger causality triggers complex networks

Tian Ge[1,2]**, Yindong Cui**[1]**, Wei Lin**[1,6]**, Jürgen Kurths**[3,4,5,6]
and Chong Liu[1]

[1] Centre for Computational Systems Biology, School of Mathematical Sciences, and Shanghai Key Laboratory of Contemporary Applied Mathematics, Fudan University, Shanghai 200433, China
[2] Department of Computer Science, The University of Warwick, Coventry CV4 7AL, UK
[3] Potsdam Institute for Climate Impact Research, D-14412 Potsdam, Germany
[4] Department of Physics, Humboldt University of Berlin, D-12489 Berlin, Germany
[5] Institute for Complex Systems and Mathematical Biology, University of Aberdeen, Aberdeen AB24 3FX, UK
E-mail: wlin@fudan.edu.cn and kurths@pik-potsdam.de

Abstract. In this paper, we propose a new approach to characterize time series with noise perturbations in both the time and frequency domains by combining Granger causality and complex networks. We construct directed and weighted complex networks from time series and use representative network measures to describe their physical and topological properties. Through analyzing the typical dynamical behaviors of some physical models and the MIT-BIH[7] human electrocardiogram data sets, we show that the proposed approach is able to capture and characterize various dynamics and has much potential for analyzing real-world time series of rather short length.

[6] Authors to whom any correspondence should be addressed.

[7] Massachusetts Institute of Technology-Beth Israel Hospital.

IOP Institute of Physics   Φ DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

Contents

1. Introduction

Characterizing typical dynamics from time series with noise perturbations is a fundamental problem eliciting a great deal of attention from various communities. Some measures of dynamical systems or information relevance, including the Lyapunov exponents, the correlation dimension or entropies, have been consecutively developed and widely applied [1]. Most of them circumscribe the investigations only in the phase spaces where time series evolve dynamically. In recent years, techniques from different realms have been proposed in [2–7], where the characterization of time series is converted from the phase space to the complex network perspective. A network is a mathematical representation of a real-world complex system and is defined by a collection of vertices and links between pairs of nodes. Therefore some representative measures describing the topological properties of a complex network can be used. A basic step in this approach is to reconstruct a complex network from a time series, i.e. to estimate connectivity weights. The linear correlation coefficient and phase space distance are commonly used to compute the connectivity weights of a complex network [2]; nonetheless, the weights of the constructed complex network are usually sensitive to the outliers existent in the time series, and are even endowed with spurious correlations. More importantly, the existing techniques [2, 5–7] cannot extract information on the time series in the frequency domain and are not applicable to non-stationary time series.

    To somewhat relax these limitations, this paper proposes an approach for integrating the Granger causality (GC) and complex networks into characterizing time series with noise perturbations. In particular, the connectivity weights of a causal complex network (Cau-CN) are obtained by computing the GC between sub-time series of a given time series.

    The GC is based on precedence and predictability. It states that given two times series x and y, if the inclusion of the past history of y helps to predict the future states of x in some plausible statistical sense, then y is a cause of x in the Granger sense. Recently, the GC, as well as its generalizations, has become one of the most powerful and valuable tools in discovering connections or causalities between different elements in various complex systems, including biological systems and climatic processes [8–11]. On the one hand, its time-domain definition leads to causal relations, not only weighted but also directed, which provides further details of the topological perspective of the interactions among those elements. On the other hand, its frequency-domain decomposition, meanwhile, makes the interpretation more informative

because those interactions even in different frequency bands could be clearly figured out instead of a single number [12]. Therefore, all these important advantages of the GC allow us to have a deeper insight into the characterization of the dynamics of the time series through the established Cau-CNs. More specifically, we will use three frequently used measures, the average path length (APL), the average vertex strength (AVS) and the average clustering coefficient (ACC) [13, 14], computed for the so constructed Cau-CNs. Interestingly, these measures in both the time and frequency domains are verified to be closely connected to typical dynamical systems. In particular, the APL and ACC show more potential in distinguishing dynamical behaviors of given time series in all frequencies. In addition, to exemplify the broad perspective of the proposed approach in real applications, this paper investigates human electrocardiogram (ECG) recordings from the MIT-BIH data sets.

2. Methods

In this section, we describe the method to construct Cau-CNs by computing the GC in the context of the linear autoregressive (AR) model in both the time and frequency domains and utilize a static as well as a dynamic approach to characterizing dynamics of time series through the network to be established.

2.1. Causal complex network construction

Given a time series consisting of n observations, denoted by $\{x[q]\}_{q=1}^{n}$, the time series-induced Cau-CN is designed as follows. Firstly, we divide the given time series equally into m pieces, which are called sub-time series, denoted by

$$\left\{ x_j(t) = x[t + (j-1)\lfloor n/m \rfloor] \right\}$$

with $j = 1, \ldots, m$ and $t = 1, \ldots, \lfloor n/m \rfloor$, where $\lfloor \cdot \rfloor$ is the floor function. Later, we will discuss the impact of the piece number as well as the length of each sub-time series. Secondly, we suppose that each sub-time series corresponds to one vertex in the finally obtained Cau-CN, denoted by $(\mathcal{V}, \mathcal{W})_{\mathcal{J}}$. Here, $\mathcal{V}$ is a set consisting of all m vertices in the network, and $\mathcal{W} = \{w_{ij}\}_{m \times m}$ is a directed weight matrix, in which each of its elements describes the GC from the vertices i to j. The subscript $\mathcal{J}$ represents an index indicating the network constructed in the time or frequency domains. We include in appendix the exact definitions of GC in both the time and the frequency domains. In particular, when $\mathcal{J} = T$, a Cau-CN $(\mathcal{V}, \mathcal{W})_T$ in the time domain is constructed, namely each weight connectivity w_{ij} of the network is computed through the formula of the causal influence $F_{i \to j}$ (see equation (A.5)); meanwhile, $\mathcal{J} = f$ theoretically invites an infinite number of networks $(\mathcal{V}, \mathcal{W})_f$ in the frequency domain, namely the weight connectivities with frequency f are obtained by the causal influence $I_{i \to j}(f)$ (see equation (A.6)), where the frequency $f \in [0, +\infty)$.

2.2. Network measures—a static characterization

Once the Cau-CN $(\mathcal{V}, \mathcal{W})_{\mathcal{J}}$ is constructed in either the time or frequency domains, not only the distributions of the weights but also the three aforementioned measures are calculated.

1. AVS describes the average of the strengths for all the network vertices [13, 14]. The strength of vertex i is defined as $s_i^{\text{tot}} = s_i^{\text{in}} + s_i^{\text{out}}$, where the in-strength $s_i^{\text{in}} = \sum_{j \neq i} w_{ji}$ and

the out-strength $s_i^{\mathrm{out}} = \sum_{j \neq i} w_{ij}$. Also the degree of vertex i is defined as $d_i^{\mathrm{tot}} = d_i^{\mathrm{in}} + d_i^{\mathrm{out}}$, where the in-degree $d_i^{\mathrm{in}} = \sum_{j \neq i} a_{ji}$, the out-degree $d_i^{\mathrm{out}} = \sum_{j \neq i} a_{ij}$ and a_{ij} is the element of the adjacency matrix A corresponding to the network $(\mathcal{V}, \mathcal{W})_{\mathcal{J}}$.

2. APL describes the average of the shortest geodesic path lengths between all pairs of network vertices, where the shortest geodesic path refers to the path between two vertices with the sum of the weights of its constituent edges minimized [13, 14].

3. ACC represents the average fraction of the connections between the topological neighbors of a vertex with respect to the maximum possible for all network vertices [13, 14]. In particular, the clustering coefficient (CC) of vertex i is defined as

$$C_i = \frac{[\mathcal{W}^{[1/3]} + (\mathcal{W}^T)^{[1/3]}]_{ii}^3}{2[d_i^{\mathrm{tot}}(d_i^{\mathrm{tot}} - 1) - 2d_i^{\leftrightarrow}]},$$

where $\mathcal{W}^{[1/3]} = \{w_{ij}^{1/3}\}$ and $d_i^{\leftrightarrow} = \sum_{j \neq i} a_{ij} a_{ji}$.

It should be noted that, in the following discussion, we use only the above three measures to characterize the constructed networks in both domains. However, there are some other typical measures that are frequently used in the characterization of complex networks [13, 14]; they have been included in our ongoing works.

2.3. Artificial neural networks—a dynamic characterization

In addition to the above static measures, we design another distinguishing method by observing the dynamics of a discrete-time artificial neural network (DANN) endowed with the weighted connection matrix obtained above. The DANN is a mathematical model inspired by the structure and functional aspects of biological neural networks which defines a discrete map to describe the evolution of the states of neurons over time. Concretely, these neurons are connected and we endow the weighted connection matrix $\{w_{ij}\}$ of a DANN:

$$N_i[s+1] = kN_i[s] + \sum_{j=1}^{m} w_{ij} M_j[s], \quad s = 0, 1, 2, \ldots$$

with a causal network by applying the procedure of section 2.1. Here, $N_i[s]$ and $M_i[s] = 1/\left[1 + e^{-N_i[s]/\epsilon}\right]$ are, respectively, the internal state and the output of neuron i at the sth iteration, and $i = 1, \ldots, m$. The damping factor and the steepness parameter are taken as $k = -0.9$ and $\epsilon = 3.3 \times 10^{-3}$, which are the typical values used for simulating a DANN [15]. The dynamics of the time series can then be characterized by the evolution of the DANN over time.

All these characterization methods are illustrated and validated in the following examples.

3. Examples: from toy models to real-world data

Example 1 (Non-chaos versus chaos). Consider a noised-perturbed time series:

$$y[q] = x[q] + \varepsilon[q], \quad q = 1, \ldots, 2000,$$

where $\varepsilon[q]$ is a zero mean independent Gaussian noise with variance 0.01, and each $x[q]$ is generated by the standard Logistic model:

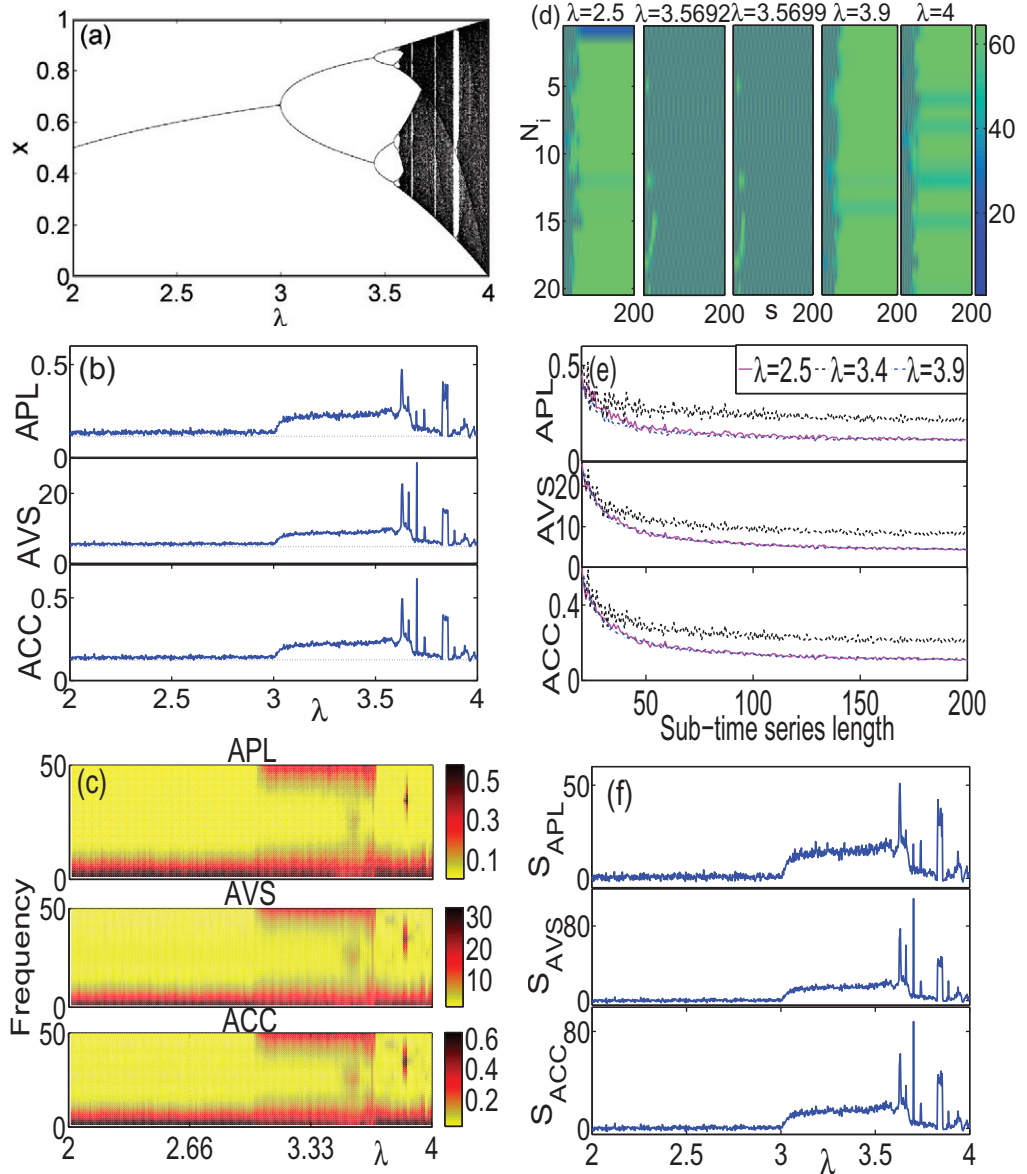$$x[q+1] = \lambda x[q]\{1 - x[q]\}, \quad q = 1, \ldots, 2000.$$

Figure 1. The Logistic model. (a) A bifurcation diagram as $\lambda \in [2, 4]$. Three measures versus λ are plotted in the time domain (b) and in the frequency domain (c). (d) Different patterns are produced by the DANNs with different casual networks. (e) Variations of the measures with the length of the sub-time series in the time domain. (f) Standard significance tests for the measures in the time domain for each λ and correspondingly 100 surrogates.

Here, with an increase of the parameter λ in the interval $[2, 4]$, the dynamics of the Logistic model show qualitative changes from stable fixed points to cascades of periodic orbits and chaos, as shown in figure 1. For each λ and, correspondingly, the noise-perturbed series $y[q]$ that contains 20 sub-time series, we establish Cau-CNs by the approach we proposed above.

According to the definition of the GC, the inclusion of the knowledge about one random sub-time series cannot significantly reduce the variance of the prediction error of another

random one. Nevertheless, the sub-time series in the periodic series, with stronger mutual causal relations, do bring about the reduction. It is thus expected that the variation of the APL, AVS and ACC of the causal network in the time domain is closely connected to the dynamics of the Logistic model with different λ. As shown in figure 1, when $\lambda \in [2, 3]$ that corresponds to the appearance of stable fixed points, $y[q]$ that consists of stable fixed point dynamics and noise perturbation is seen as a random series, so that the weak causal internal relations between its sub-time series naturally result in small values of the network measures. The measures become larger when λ is taken from the region where stable periodic dynamics appear and again become smaller as λ is in the parameter region of chaos. Note also that, although the measures for the chaotic dynamics are relatively small, most of them are still larger than those for the random series as $\lambda \in [2, 3]$. This somewhat implies the difference between chaos and randomness from a casual network point of view. In addition, the peaks of the measurement curves in figure 1 correspond to the islands of stable periodic dynamics embedded in the chaotic region, which further illustrates the reliability of the proposed approach in differentiating between chaos and non-chaotic dynamics. To be candid, this kind of differentiation can be analogously done by some existing techniques [5–7]; however, the proposed approach is further capable of showing the frequency bands where the differences appear between the causal networks corresponding to different dynamics. As shown in figure 1, the three measures become larger in both lower and some higher frequency bands when λ takes the value by which the Logistic model has periodic dynamics. This observation indeed implies some concrete frequency bands at which causal relations among the sub-time series of the periodic time series are stronger.

The time series generated from the Logistic model with different parameters are also characterized by the dynamics of their corresponding DANNs. As interestingly shown in figure 1, all the states of the neurons converge to the fixed points as λ corresponds to random or chaotic behavior in the noise-perturbed series $y[q]$; however, the states converge to periodic or quasi-periodic orbits, forming thin stripe patterns, as λ takes a value by which the Logistic model generates periodic dynamics. Hence, DANNs, when endowed with Cau-CNs, can preserve and even be used to differentiate some dynamical properties of time series for rather short data sets.

To validate the robustness of the proposed approach with regard to the length of the sub-time series, the measures for the networks in the time domain are, respectively, plotted versus length in figure 1, where the curves become flatter as the length increases. This observation is consistent with the result, reported in [16], on the length requirement (at least 100 points) of the time series for computing the GC. Also this illustrates that our method is applicable for analyzing a rather short data set.

In addition, random phase surrogates and standard significance tests [17] are utilized to further validate the proposed approach. Specifically, for each λ, the generated time series are randomly permuted for 100 times. Surrogates of the measure APL are then computed based on the Cau-CNs constructed from these permuted time series. We denote the mean and standard deviation (SD) of these surrogates by m_{APL} and σ_{APL}, respectively. The statistic S_{APL} is defined as $S_{APL} = (APL - m_{APL})/\sigma_{APL}$, which measures the deviation of APL from its surrogates. S_{AVS} and S_{ACC} are defined similarly. It can be seen from figure 1 that when λ corresponds to periodic dynamics, the three network measures are significantly different from their surrogates, while when λ corresponds to random or most of the chaotic time series, the difference, although not very outstanding, can still be identified. If we need to identify the difference significantly between random and chaotic dynamics, then in addition to the AR model some nonlinear models should also be utilized in the computation of the GC.
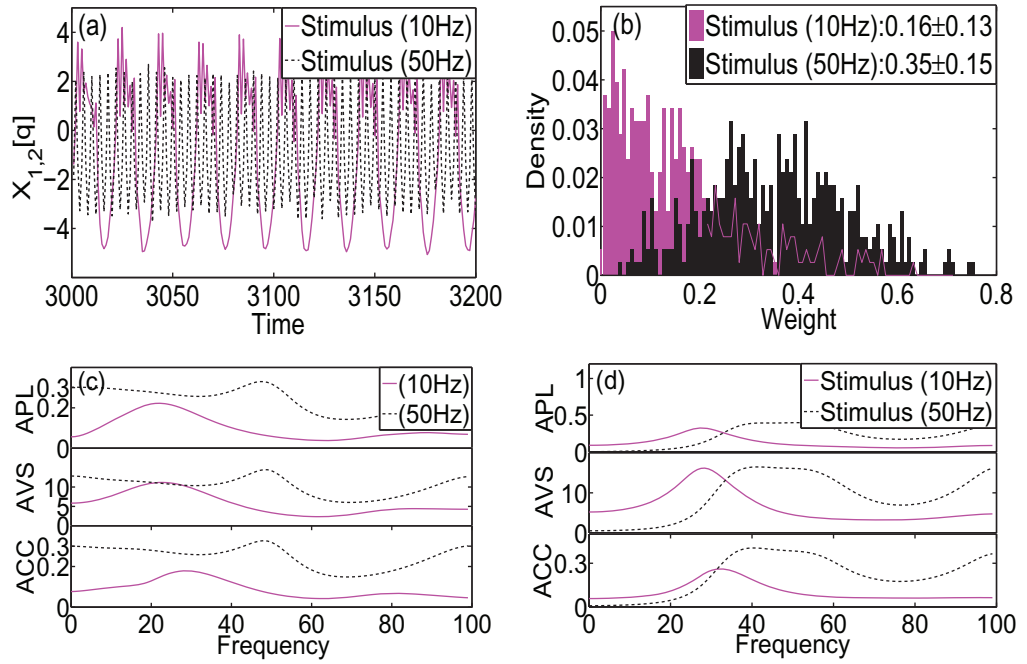
Figure 2. The Rulkov model with stimuli. (a) Two time series with spikes and bursts. (b) Distributions of the matrix weights in the time domain. Measurements in the frequency domain for the original two time series (c) and for the two time series after the first-order difference (d).

Example 2 (Low-frequency versus high-frequency). Next we consider neural dynamics, described by the Rulkov model [18, 19], with sinusoidal stimuli in two frequencies. We generate two time series $\{x_i[q]\}$ $(i = 1, 2)$, corresponding to the membrane potential of the neuron, as follows:

$$x_i[q + 1] = \frac{\alpha}{1 + x_i^2[q]} + y_i[q] + \eta \sin\left(\frac{2\pi f_i^* q}{\mathrm{sr}}\right) + \zeta \varepsilon_i[q],$$

$$y_i[q + 1] = y_i[q] - \sigma x_i[q] - \beta.$$

Here, $x_i[1] = y_i[1] = 0.1$, the two intrinsic frequencies for the sinusoidal stimuli are set as $f_1^* = 10\,\mathrm{Hz}$ and $f_2^* = 50\,\mathrm{Hz}$, and the sampling rate, $\mathrm{sr} = 200\,\mathrm{Hz}$. Each $\varepsilon_i[q]$ is a zero mean uncorrelated Gaussian noise with variance 1.0, and the strengths of the stimuli and the noise are $\eta = 3.0$ and $\zeta = 0.1$, respectively. The other parameters are taken as $\alpha = 4.2$ and $\sigma = \beta = 0.001$, so that the model exhibits chaotic bursting dynamics without any stimulus. Then, we use the obtained time series (see figure 2) to construct the Cau-CNs, respectively. Here, each time series contains 5000 points, and the last 2000 points, divided into 20 sub-time series, are considered.

As shown in figure 2, a clear difference between both distributions of the matrix weights of the constructed networks in the time domain can be observed. For the stimulus of lower frequency, most of the weights take smaller values because the chaos property of the bursting time series keeps still; however, more regular spikes due to the stimulus of high frequency emerge, yielding a right shift of the distribution. In addition, the three measures for the constructed networks in the frequency-domain approach maxima around some specific
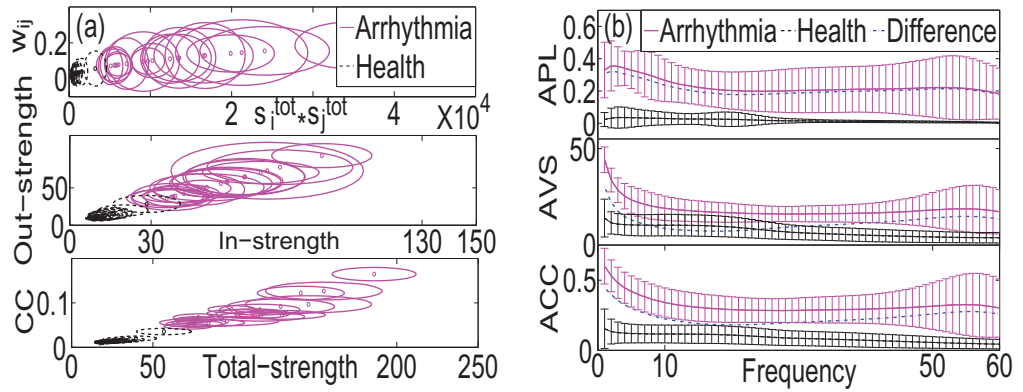
Figure 3. Results of the MIT-BIH ECG data sets. (a) The weight w_{ij} versus $s_i^{tot} \times s_j^{tot}$ (upper panel), out-strength versus in-strength (middle panel) and clustering coefficient versus total strength (lower panel) for each subject. The center of each ellipse represents the average of the measures within each subject, and the lengths of the ellipse's semi-major and semi-minor axes represent the SDs in the respective directions. (b) Averages and SDs of the measures across subjects from the two data sets, respectively, in the frequency domain. Differences between the averages of the measures for the two data sets are marked by the dash-dotted (blue) lines.

frequencies (see figure 2). These frequencies clearly correspond to either the intrinsic frequency of the chaotic bursting or the high frequency of the stimuli. More importantly, there is no overlap between the curves of APL and ACC across different frequencies in figure 2. In fact, these two quantities, rather than the AVS, reflect the global property of the underlying Cau-CNs. They therefore show more potential in differentiating the different time series in all frequencies.

Besides, after the first-order difference of the time series, that is, $\hat{x}_i[q] = x_i[q+1] - x_i[q]$, the maxima positions of the measures do not change much (figure 2). This implies that the proposed approach could be suitable for characterizing some non-stationary time series, since the difference procedure is commonly used to render time series stationary and thus meet the requirements of classical Granger causal analysis.

Example 3 (Health versus arrhythmia). Now we analyze real-world data. Two MIT-BIH ECG data sets[8] are used in the analysis: one is from the Normal Sinus Rhythm Database (NSRD) and the other is from the Arrhythmia Database (AD). We select the first 18 subjects from each of the two data sets. ECG recorded the electrical activity of the heart and produced a time series for each subject covering about 20 min. 500 sub-time series are used to construct the corresponding Cau-CNs based on the above-proposed approach in the time domain. Here, after selecting a sub-time series that contains 500 points, we set the next R-peak as the start point of the following sub-time series and calculate the three network quantities mentioned above. For the NSRD, the means and SDs of the APL, AVS and ACC are 0.00067 ± 0.00049, 26.7 ± 9.5 and 0.015 ± 0.006, respectively, whereas those for the AD are 0.028 ± 0.020, 110.3 ± 32.5 and 0.081 ± 0.030, respectively. More striking differences are shown in figure 3, which characterizes the in- and out-strengths and CCs of the vertices in the constructed Cau-CNs. All these

---

[8] All the data sets are available at http://www.physionet.org/physiobank/database.

differences suggest that stronger causal relations exist between the sub-time series for the arrhythmia subjects, but weaker causal relations for the normal sinus rhythm subjects. This also indicates that the heart rhythm of healthy people adapts more sensitively to internal or environmental fluctuations. These somewhat random fluctuations reasonably suppress the relations quantified by the GC.

To more profoundly study the difference between both data sets, we also establish and analyze the causal networks in the frequency domain. Here, we use 20 sub-time series, each of which contains 200 points. As shown in figure 3, the averages of the three quantities for the two data sets are different in all frequency bands. More importantly, the proposed approach enables us to figure out the frequency band (1–2 Hz) where the differences between the healthy subjects and arrhythmia patients are most prominent. This band interestingly corresponds to a human's normal heart rate, approximately 60–120 beats per minute. Besides, in figure 3, more overlaps of the SDs of the two datasets are observed for the AVS rather than the other two measures. The reason for this has been stressed in example 2.

4. Conclusion

Altogether, in this paper we have proposed a new approach to characterizing time series by combining the GC and the concept of complex networks. Directed and weighted complex networks constructed by the GC contain more details of the topological perspective of the interactions among sub-time series, and provide more possibilities to capture the various dynamics of the time series when compared with undirected and unweighted networks. More importantly, decomposition of the GC in the frequency domain enables the construction of frequency-specific complex networks, making it possible to characterize the time series at a particular frequency or frequency band. This is particularly useful in analyzing real-world time series when different frequency bands convey different signals and have distinct functions (e.g. local field potential or electroencephalography data) or when some frequencies are contaminated by noise and contain little information (e.g. functional magnetic-resonance-imaging data). Furthermore, due to the robustness of GC, the proposed approach can be also applied to non-stationary time series. By computing three representative network characteristics of the established causal networks, the proposed approach has been applied to characterize the time series produced by toy dynamical models and the recordings from two MIT-BIH data sets, in both the time and frequency domains. The results show the potential of our approach to capture various dynamics and to discriminate healthy subjects from arrhythmia patients.

Future works involve applying the proposed approach to analyze evolving time series by using a sliding window and constructing Cau-CNs, which may thus be applicable to rather short data sets. Moreover, since only the AR model is used for computing the GC, integrating the other nonlinear models [20, 21] as well as more measures of complex networks into the present approach might provide a more accurate characterization, which also deserves further investigation.

Appendix. The Granger causality in the time and frequency domains

We give a brief review of the GC in the context of the linear AR model in both the time and frequency domains. Suppose x_1 and x_2 to be two zero-mean signals whose time observations are denoted by $x_1(t)$ and $x_2(t)$, where $t = 1, \ldots, T$. The univariate AR models are

$$x_1(t) = \sum_{j=1}^{p} \alpha_j x_1(t-j) + \epsilon_1(t) \tag{A.1}$$

and

$$x_2(t) = \sum_{j=1}^{p} \beta_j x_2(t-j) + \epsilon_2(t), \tag{A.2}$$

where the prediction errors $\epsilon_1(t)$ and $\epsilon_2(t)$ are uncorrelated over time. The model considered by Granger assumes that $x_1(t)$ and $x_2(t)$ are suitably represented by the bivariate model:

$$x_1(t) = \sum_{j=1}^{p} a_{11,j} x_1(t-j) + \sum_{j=1}^{p} a_{12,j} x_2(t-j) + \epsilon_{1|2}(t) \tag{A.3}$$

and

$$x_2(t) = \sum_{j=1}^{p} a_{21,j} x_1(t-j) + \sum_{j=1}^{p} a_{22,j} x_2(t-j) + \epsilon_{2|1}(t). \tag{A.4}$$

Here, p is the order of the AR model, determined by the standard Akaike information criterion in this paper, and ϵ's represent the prediction errors. If the variances of the prediction errors satisfy $\mathrm{var}[\epsilon_{1|2}(t)] < \mathrm{var}[\epsilon_1(t)]$, x_2 is said to have a causal influence on x_1. Analogously, if $\mathrm{var}[\epsilon_{2|1}(t)] < \mathrm{var}[\epsilon_2(t)]$, then there is a causal influence from x_1 to x_2. These causal influences are characterized in the time domain by

$$F_{j \to i} = \ln \left\{ \mathrm{var}[\epsilon_i(t)] / \mathrm{var}[\epsilon_{i|j}(t)] \right\}, \tag{A.5}$$

where $i = 1, 2$ and correspondingly $j = 2, 1$. Moreover, Geweke [12] proposed a spectral decomposition of the GC from the time domain to the frequency domain. With a lag operator $\mathcal{L} : \mathcal{L}x(t) = x(t-1)$, the above bivariate model in a matrix form becomes

$$A(\mathcal{L})[x_1(t), x_2(t)]^{\mathrm{T}} = [\epsilon_{1|2}, \epsilon_{2|1}]^{\mathrm{T}},$$

where $A(\mathcal{L}) = \{a_{ij}(\mathcal{L})\}_{2 \times 2}$ and $a_{ij}(\mathcal{L}) = \sum_{k=0}^{p} a_{ij,k} \mathcal{L}^k$ with $a_{ij,0} = \delta_{ij}$. The covariance matrix of the noise terms is $\Sigma = \{\Sigma_{ij}\}_{2 \times 2}$, where $\Sigma_{12} = \Sigma_{21} = \mathrm{cov}[\epsilon_{1|2}, \epsilon_{2|1}]$ and $\Sigma_{ii} = \mathrm{var}[\epsilon_{i|j}]$ for $i \neq j$. Thus, the discrete Fourier transformation of the model in the matrix form yields the spectral equivalent model:

$$A(f)[x_1(f), x_2(f)]^{\mathrm{T}} = [E_1(f), E_2(f)]^{\mathrm{T}},$$

where

$$A(f) = \{a_{ij}(f)\}_{2 \times 2}, \quad a_{ii}(f) = 1 - \sum_{k=1}^{p} a_{ii,k}\, \mathrm{e}^{-2\mathrm{i}\pi k f}, \quad a_{ij}(f) = -\sum_{k=1}^{p} a_{ij,k}\, \mathrm{e}^{-2\mathrm{i}\pi k f}$$

for $i \neq j$. Hence, setting the transfer function matrix $H(f) = \{a_{ij}(f)\}_{2\times 2}^{-1}$ leads the spectral model above to

$$[x_1(f), x_2(f)]^{\mathrm{T}} = H(f)[E_1(f), E_2(f)]^{\mathrm{T}}.$$

Then, the spectral density matrix $S(f)$ is given by $S(f) = H(f)\Sigma H^*(f)$. To eliminate the cross terms in $S(f)$, Geweke's normalization technique [12] is further adopted. Consequently, the GC from x_j to x_i at the frequency f is

$$I_{j\to i}(f) = \ln\left\{ S_{ii}(f) \Big/ \left[S_{ii}(f) - \left(\Sigma_{jj} - \Sigma_{ij}^2/\Sigma_{ii} \right) |H_{ij}(f)|^2 \right] \right\}. \tag{A.6}$$

References

[1] Abarbanel H D I, Brown R, Sidorowich J J and Tsimring L Sh 1993 *Rev. Mod. Phys.* **65** 1331
[2] Zhang J and Small M 2006 *Phys. Rev. Lett.* **96** 238701
[3] Zhang J and Luo X 2007 *Phys. Rev.* E **75** 016218
[4] Xu X, Zhang J and Small M 2008 *Proc. Natl Acad. Sci. USA* **105** 19601
[5] Marwan N *et al* 2009 *Phys. Lett.* A **373** 4246
[6] Donner R V *et al* 2010 *New J. Phys.* **12** 033025
[7] Donner R V *et al* 2010 *Phys. Rev.* E **81** 015101
[8] Granger C W J 1969 *Econometrica* **37** 424
[9] Gourevitch B *et al* 2006 *Biol. Cybern.* **95** 349
[10] Friston K 2009 *PLoS Biol.* **7** e1000033
[11] Mokhov I I *et al* 2011 *Geophys. Res. Lett.* **38** L00F04
[12] Geweke J 1984 *J. Am. Stat. Assoc.* **79** 907
[13] Albert R and Barabasi A L 2002 *Rev. Mod. Phys.* **74** 47
[14] Fagiolo G 2007 *Phys. Rev.* E **76** 026107
[15] Lin W and Chen G 2009 *IEEE Trans. Neural Netw.* **20** 1340
[16] Zou C and Feng J 2009 *BMC Bioinf.* **10** 122
[17] Theiler J *et al* 1992 *Physica* D **58** 77
[18] Rulkov N F 2001 *Phys. Rev. Lett.* **86** 183
[19] Rulkov N F 2002 *Phys. Rev.* E **65** 041922
[20] Dhamala M *et al* 2008 *Phys. Rev. Lett.* **100** 018701
[21] Marinazzo D *et al* 2008 *Phys. Rev. Lett.* **100** 144103