Commun Nonlinear Sci Numer Simulat 51 (2017) 78–88

Contents lists available at ScienceDirect

Commun Nonlinear Sci Numer Simulat

journal homepage: www.elsevier.com/locate/cnsns

Research paper

Symbolic phase transfer entropy method and its application

Ningning Zhang, Aijing Lin*, Pengjian Shang

School of Science, Beijing Jiaotong University, Beijing 100044, PR China

ARTICLE INFO

Article history:
Received 24 May 2016
Revised 30 January 2017
Accepted 12 March 2017
Available online 21 March 2017

Keywords:
Symbolic phase transfer entropy (SPTE)
Symbolic transfer entropy (STE)
Information flow
Stocks

ABSTRACT

In this paper, we introduce symbolic phase transfer entropy (SPTE) to infer the direction and strength of information flow among systems. The advantages of the proposed method are investigated by simulations on synthetic signals and real-world data. We demonstrate that symbolic phase transfer entropy is a robust and efficient tool to infer the information flow between complex systems. Based on the study of the synthetic data, we find a significant advantage of SPTE is its reduced sensitivity to noise. In addition, SPTE requires less amount of data than symbolic transfer entropy(STE). We analyze the direction and strength of information flow between six stock markets during the period from 2006 to 2016. The results indicate that the information flow among stocks varies over different periods. We also find that the interaction network pattern among stocks undergoes hierarchial reorganization with transition from one period to another. It is shown that the clusters are mainly classified according to period, and then by region. The stocks during the same time period are shown to drop into the same cluster.

1. Introduction

It is a general topic of common interest to understand the information flow between nonlinear systems [1–6]. More specifically, it includes two important indicators: the direction and strength. The transfer entropy is a commonly used tool of nonlinear systems. The transfer entropy which measures the information flow between nonlinear systems has been recently introduced by Schreiber [7]. A. Kaiser and T. Schreiber [8], F. Verdes [9] and Y. Kuniyoshi [10] point out that all kinds of technologies have been proposed to estimate the transfer entropy from the observed values [11–16]. However, most of the technologies highly require for data, need to adjust parameters, and are sensitive to noise. The shortcomings hinder the wide application of transfer entropy in various fields.

In 2008, M. Staniek and K.K. Lehnertz numerically illustrate that symbolic transfer entropy in quantifying dominating direction of information flow between time series from coupling system of the same and different structure [17], is a powerful and fast calculation method. These features make the symbolic transfer entropy a promising method in data analysis. Through investigation and study on the numerical and brain electrical activity time series analysis, it is found that asymmetric dependence between the same and different structure, but not fully synchronous coupling system.

M. Lobier, F. Siebenhhner, S. Palva and J.M. Palva propose phase transfer entropy (Phase TE) as a measure to infer the direct relationship between neural oscillations [18]. Phase TE is a new, phase-based measure of directed connectivity. Moreover, they find that Phase TE detects the strength and direction of the connections even in the presence of vast amounts of noise and linear mixed. Phase TE is robust to nuisance parameters and more efficient computationally. Phase TE

* Corresponding author.
E-mail address: ajlin@bjtu.edu.cn (A. Lin).

http://dx.doi.org/10.1016/j.cnsns.2017.03.011
1007-5704/

detects connectivity between time series across a range of analysis lag. Finally, because Phase TE is similar to phase-specific functional connectivity metrics, it could identify information flow of limited frequency band. Those indicate that Phase TE is suitable for estimates of the direct connections based on the Phase.

Therefore, considering of faults and the bottleneck problem in the study of transfer entropy method, this paper proposes the improved method of transfer entropy to overcome the limitations of existing methods, which is novel in detection of the information flow among nonlinear systems. This paper intends to study symbolic phase transfer entropy based on the transfer entropy method, which is used to quantify the direction and strength of information flow among systems and applied to the experiment analysis for synthetic data and real-world data. This paper mainly uses symbolic phase transfer entropy to study nonlinear causality among the stocks.

In this paper, we quantify symbolic phase transfer entropy among DAX, FTSE, S&P500, NAS, ShangZheng and ShenCheng stock closing prices. In order to accurately test the validity of proposed method, we consider time series generated by AR model, ARFIMA model and Henon map. The proposed SPTE method is shown to cater for detecting the direction and strength of information flow among systems.

The rest of the paper is organized as follows. In Section 2, we introduce symbolic transfer entropy method and symbolic phase transfer entropy method in detail. In Section 3, we provide artificial data and real stock data. Analysis and results for data are shown in Section 4. The conclusions are presented at the last section.

2. Methodology

2.1. Symbolic transfer entropy

The method of symbolic transfer entropy is based on the phase space reconstruction and associated with the definition of permutation entropy [19]. For an arbitrary i, $X(i)=\{x(i),x(i+1),\dots,x(i+(m-1)l)\}$ got by phase space reconstruction are arranged in an ascending order $\{x(i+(k_{i1}-1)l)\le x(i+(k_{i2}-1)l)\le\dots\le x(i+(k_{im}-1)l)\}$, where l is the time delay, and m denotes the embedding dimension. In case of equal values, for example, while $x(i+(k_{i1}-1)l)=x(i+(k_{i2}-1)l)$ we write $x(i+(k_{i1}-1)l)\le x(i+(k_{i2}-1)l)$ if $k_{i1}\le k_{i2}$, therefore we insure that every X_i is uniquely mapped onto one of the $m!$ possible permutations. A symbol is thus defined as $\hat{x}_i=(k_{i1},k_{i2},\dots,k_{im})$, and with the relative frequency of symbols we estimate joint and conditional probabilities of the sequence of permutation indices. Given two time sequences $\{\hat{x}_i\}$ and $\{\hat{y}_i\}$, we define symbolic transfer entropy as

$$\begin{aligned}T^S_{Y,X} &= -\sum p(\hat{x}_{i+\delta},\hat{x}_i,\hat{y}_i)\log p(\hat{x}_{i+\delta}\mid\hat{x}_i)\\ &\quad+\sum p(\hat{x}_{i+\delta},\hat{x}_i,\hat{y}_i)\log p(\hat{x}_{i+\delta}\mid\hat{x}_i,\hat{y}_i)\\ &=\sum p(\hat{x}_{i+\delta},\hat{x}_i,\hat{y}_i)\log\frac{p(\hat{x}_{i+\delta}\mid\hat{x}_i,\hat{y}_i)}{p(\hat{x}_i\mid\hat{y}_i)}\end{aligned}\tag{1}$$

where

$$p(\hat{x}_{i+\delta}\mid\hat{x}_i,\hat{y}_i)=\frac{p(\hat{x}_{i+\delta},\hat{x}_i,\hat{y}_i)}{p(\hat{x}_i,\hat{y}_i)}\tag{2}$$

$$p(\hat{x}_{i+\delta}\mid\hat{x}_i)=\frac{p(\hat{x}_{i+\delta},\hat{x}_i)}{p(\hat{x}_i)}\tag{3}$$

where the sum runs over all symbols and δ denotes a time step. $T^S=T^S_{Y,X}-T^S_{X,Y}$ qualifies the direction of information flow between X and Y. If the value of T^S is positive, the direction of information flow is from Y to X. On the contrary, the direction of information flow is from X to Y. The value represents the strength of information flow.

2.2. Symbolic phase transfer entropy

The symbolic phase transfer entropy is symbolic transfer entropy based on phase. That is to say, before we deal with the ordinary time series, we extract phase time series from the original time series. Then we deal with the phase time series. Therefore, we define symbolic phase transfer entropy in the following steps.

Step 1: For the given time series $X(i)$ and $Y(i)$, its instantaneous phase time-series $\theta^x(t)$ and $\theta^y(t)$ is separately obtained by $X(i)$ and $Y(i)$ with Hilbert transform [20]. For the given time series x, its analytic series ζ is defined as

$$\zeta(t)=x(t)+j\hat{x}(t)=A(t)e^{j\theta(t)}\tag{4}$$

$x(t)$ with Hilbert transform gets $\hat{x}(t)$

$$\hat{x}(t)=\frac{1}{\pi}P.V.\cdot\int_{-\infty}^{+\infty}\frac{x(t)}{t-\tau}d\tau\tag{5}$$

P.V. denotes the integral of the Cauchy principal value. In Eq. (4), $A(t)$ is the instantaneous amplitude and $\theta(t)$ is the instantaneous phase of $x(t)$.

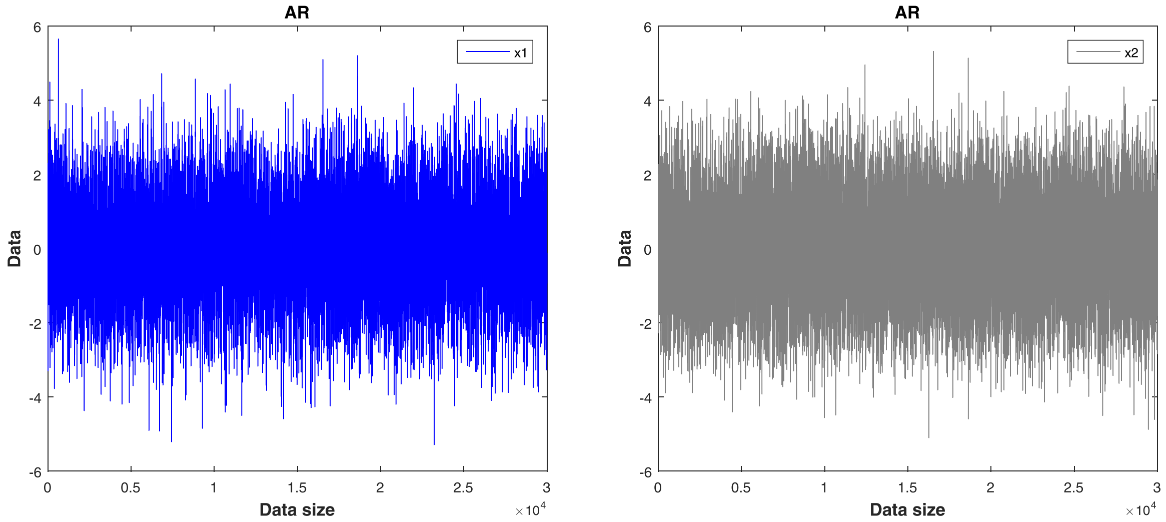

Fig. 1. The picture shows the two series $x1(i)$, $x2(i)$ of the AR model.

Table 1
The list of seven stock markets.

Num	Area	Index
1	Europe	DAX
2	Europe	FTSE
3	America	NAS
4	America	S&P500
5	China	ShangZheng
6	China	ShenCheng

Step 2: We calculate the symbolic transfer entropy of instantaneous phase time-series. Then we can get two instantaneous phase time-series $\theta^x(t)$ and $\theta^y(t)$. First, we reconstruct the phase space of $\theta^x(t)$ and $\theta^y(t)$, and then the symbolic processing is carried out. With the relative frequencies of the symbols, we estimate the joint and conditional probabilities of the sequence of permutation index. Finally, we compute the corresponding probability. The symbolic phase transfer entropy from Y to X is

$$\begin{aligned} SPTE_{Y,X} &= -\sum p(\hat{\theta}^x_{i+\delta}, \hat{\theta}^x_i, \hat{\theta}^y_i) \log p(\hat{\theta}^x_{i+\delta} \mid \hat{\theta}^x_i) \\ &\quad + \sum p(\hat{\theta}^x_{i+\delta}, \hat{\theta}^x_i, \hat{\theta}^y_i) \log p(\hat{\theta}^x_{i+\delta} \mid \hat{\theta}^x_i, \hat{\theta}^y_i) \\ &= \sum p(\hat{\theta}^x_{i+\delta}, \hat{\theta}^x_i, \hat{\theta}^y_i) \log \frac{p(\hat{\theta}^x_{i+\delta} \mid \hat{\theta}^x_i, \hat{\theta}^y_i)}{p(\hat{\theta}^x_i \mid \hat{\theta}^y_i)} \end{aligned} \tag{6}$$

where

$$p(\hat{\theta}^x_{i+\delta} \mid \hat{\theta}^x_i, \hat{\theta}^y_i) = \frac{p(\hat{\theta}^x_{i+\delta}, \hat{\theta}^x_i, \hat{\theta}^y_i)}{p(\hat{\theta}^x_{i+\delta}, \hat{\theta}^y_i)} \tag{7}$$

$$p(\hat{\theta}^x_{i+\delta} \mid \hat{\theta}^x_i) = \frac{p(\hat{\theta}^x_{i+\delta}, \hat{\theta}^x_i)}{p(\hat{\theta}^x_{i+\delta})} \tag{8}$$

3. Data

3.1. Artificial data

we simulate three well-known dynamical systems respectively generated by linear AR model [21], ARFIMA model [22–28] and Henon map to verify the validity of the proposed method. First, for the AR model, we can generate the following two signals

$$\begin{aligned} X_n &= 0.6X_{n-1} - 0.4X_{n-4} + W_n, \\ Y_n &= 0.6X_{n-1} + W'_n, \end{aligned} \tag{9}$$

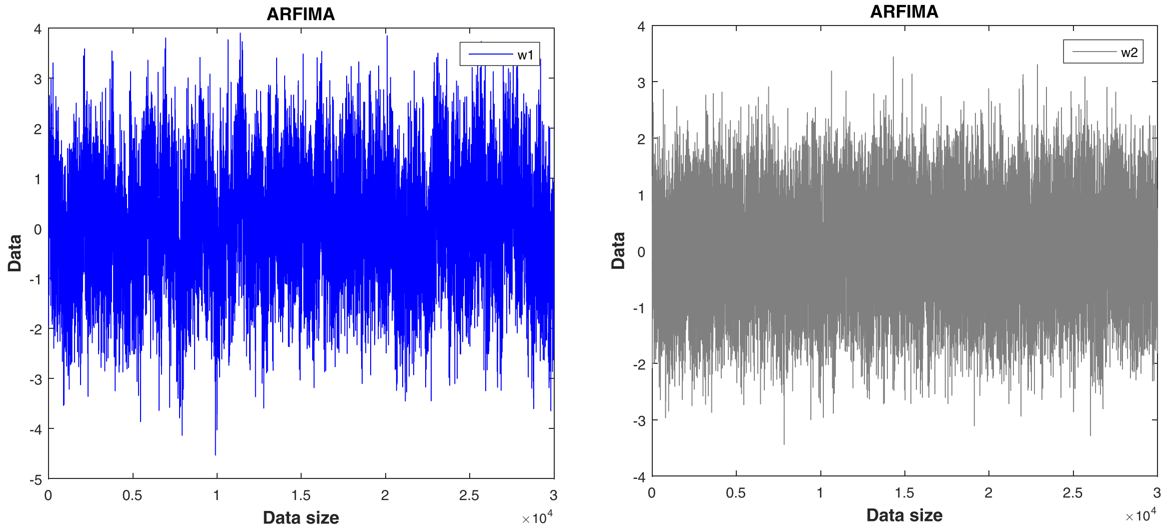

Fig. 2. The picture shows the two series $w1(i)$, $w2(i)$ of the ARFIMA model for $\alpha_1 = 0.4$, $\alpha_2 = 0.1, W = 0.8$.

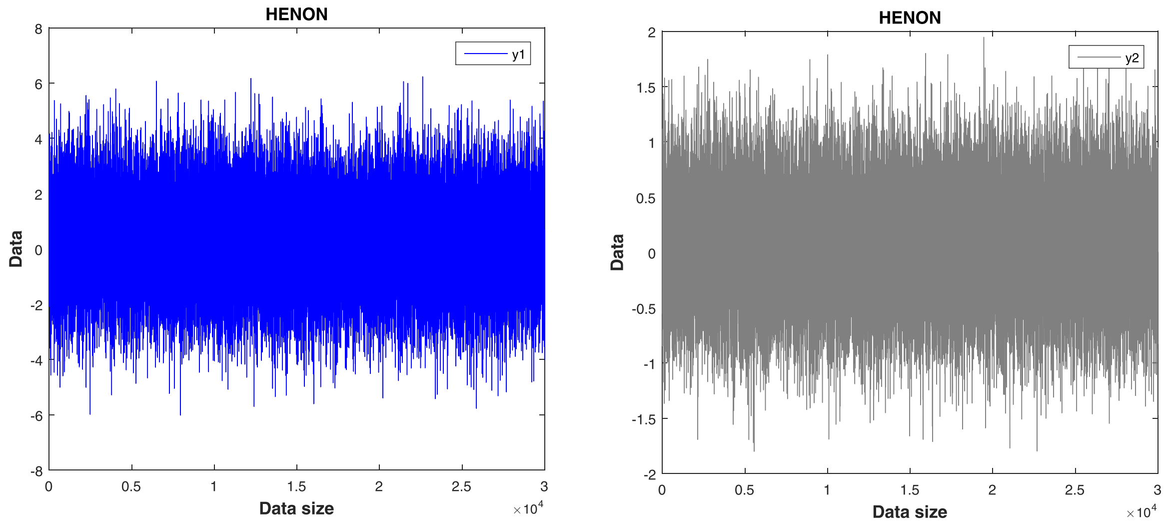

Fig. 3. The picture shows the two series $y1(i)$, $y2(i)$ of the henon map for $a = 1.3, b = 0.3$.

where W_n and W'_n are independent white Gaussian noises, and those are with zero means and unit variances. It becomes obvious that $\{\hat{y}_n\}$ is totally decided by $\{\hat{x}_n\}$. Next, for ARFIMA model, we can generate two signals by the following equation

$$
\begin{aligned}
x_i &= \left[W \sum_{n=1}^{\infty} a_n(\alpha_1) x_{i-n} + (1-W) \sum_{n=1}^{\infty} a_n(\alpha_2) x'_{i-n} \right] + A_1 \sin\left(\frac{2\pi}{T} i\right) + \eta_i \\
x'_i &= \left[(1-W) \sum_{n=1}^{\infty} a_n(\alpha_1) x_{i-n} + W \sum_{n=1}^{\infty} a_n(\alpha_2) x'_{i-n} \right] + A_2 \sin\left(\frac{2\pi}{T} i\right) + \eta'_i
\end{aligned} \tag{10}
$$

Here, η_t and η'_t indicate two Gaussian variables with zero mean and unit variance, and those are independent and identically distributed (i.i.d.). T denotes the sinusoidal period, $a_j(\alpha_m)$ are defined by Gamma function: $\alpha_m = -\frac{\Gamma(j-\alpha_m)}{\Gamma(-\alpha_m)\Gamma(1+j)}$ where $\Gamma(x)$ is the Gamma function, the range of α_m(for m = 1,2) is from 0 to 0.5, A_1 and A_2 are two amplitudes. In addition, the range of W is from 0.5 to 1. Finally, Henon map is introduced by M. Hénon as a simplified model of the Poincaré section of the Lorenz model. We can get two series as follows

$$
\begin{aligned}
x' &= 1 - ax^2 + y, \\
y' &= bx,
\end{aligned} \tag{11}
$$

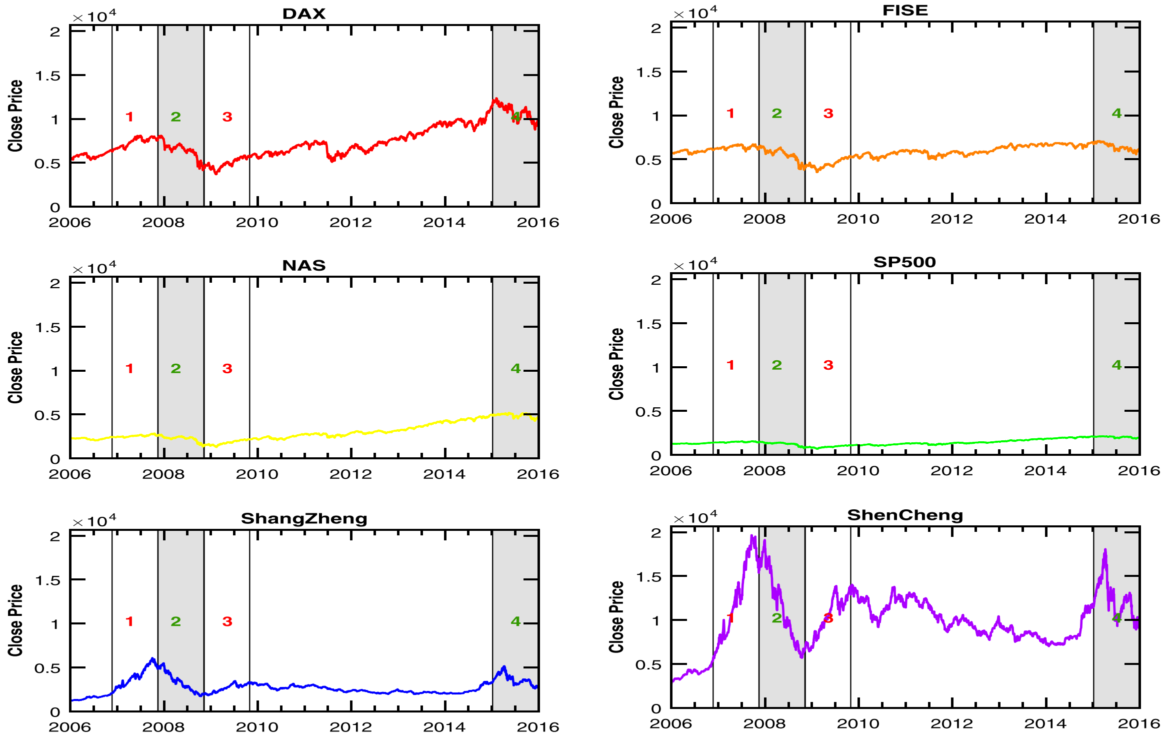

Fig. 4. Stocking closing prices of DAX, FTSE, NAS, S&P500, ShangZheng and ShenCheng. 1, 2, 3, 4 in figure respectively represent period 1, period 2, period 3 and period 4.

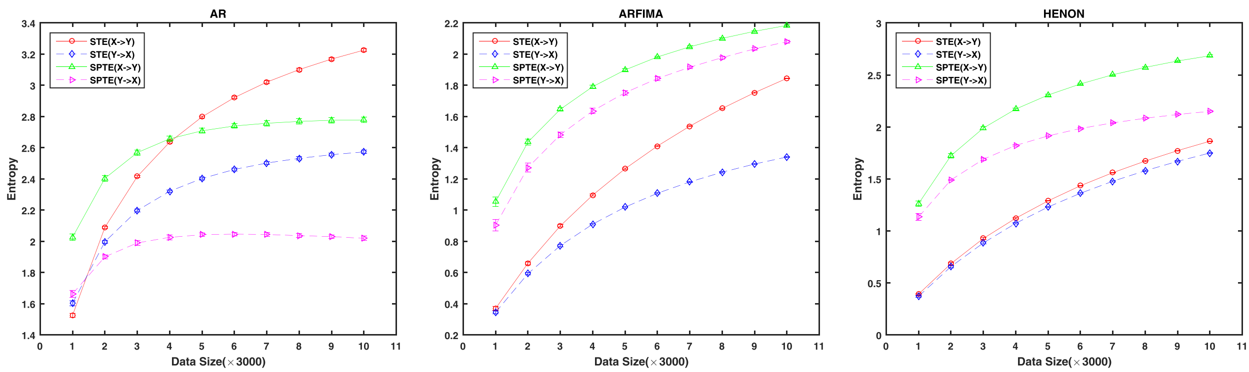

Fig. 5. Dependence between the symbolic phase transfer entropy and total length of time series. For each data size, we generate one pair of time series by AR model, ARFIMA model and Henon map. *X*-axis represents data size and *Y*-axis represents entropy computed by SPTE and STE between two signals. The curves represent an average of 20 independent realizations and error bars the standard deviation(SD).

x and y are separately initial value of x' and y'.

This paper investigates the model systems which are generated data by linear AR model, ARFIMA model and Henon map and the real stock data. First, with linear AR model, we generate two series $x1(i)$, $x2(i)$ (see Fig. 1) with 30,000 data points. Second, we perform ARFIMA procedure to get two series $w1(i)$, $w2(i)$ (see Fig. 2) with the same sample length of 30000. Then, we get two pairs of time series $y1(i)$, $y2(i)$ (see Fig. 3) with the same sample length of 30,000 based on Henon map to verify our hypothesis.

3.2. Real-world stock data

In an effort to illustrate the corresponding capabilities of the SPTE method for real-world dynamics, we also analyze six stocks of the America, Europe and China during the period from January 3, 2006 to March 11, 2016 which are listed in Table 1. We obtain data sets from the website of http://finance.yahoo.com. Due to stock markets have the different trading dates, we complement the asynchronous data with cubic spline interpolation and then reconnect the remaining parts of the

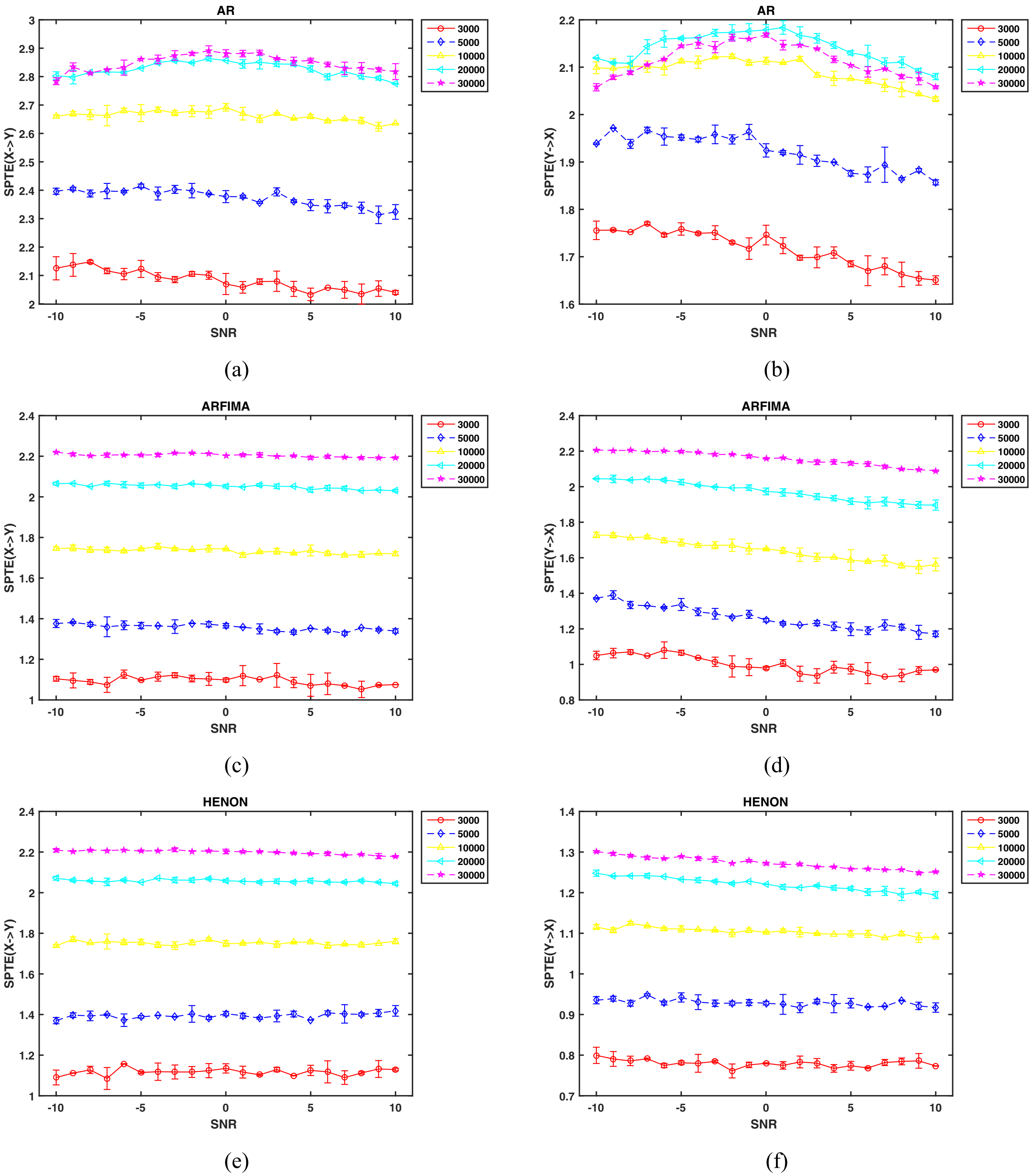

Fig. 6. The picture shows that SPTE(X-Y) and SPTE(Y-X) between two signals generated by AR model((a), (b)), ARFIMA model((c), (d)), Henon map((e), (f)) changes with the changes of the value of the SNR. And we generate 5 different lengths of signals as representatives. The curves represent an average of 20 independent realizations and error bars the standard deviation (SD).

original series to obtain the same length time series. Considering the information flow between stocks in the overall 10 years is complex, the time period under investigation in this work can be divided into alternating periods of two bullish periods and two bearish periods. The bullish periods are from December 30, 2006 to December 30, 2007 and from December 30, 2008 to December 30, 2009. The bearish periods are from December 30, 2007 to December 30, 2008 and from March 11, 2015 to March 11, 2016. The six stock data sets after the preprocessing are shown in Fig. 4.

Table 2
The values of the symbolic phase transfer entropy.

DAX	FTSE	NAS	S&P500	ShangZheng	ShenCheng
Period 1	0.6288	0.6947	0.6424	0.6475	0.6410
Period 2	0.5028	0.6653	0.6226	0.6233	0.6216
Period 3	0.4088	0.5269	0.6210	0.7151	0.6704
Period 4	0.5213	0.6689	0.6791	0.6568	0.6526
FTSE	DAX	NAS	S&P500	ShangZheng	ShenCheng
Period 1	0.5874	0.6986	0.6788	0.6834	0.6658
Period 2	0.4752	0.6999	0.6530	0.6480	0.6231
Period 3	0.4466	0.5897	0.6564	0.7313	0.6684
Period 4	0.6165	0.5850	0.6330	0.6407	0.6537
NAS	DAX	FTSE	S&P500	ShangZheng	ShenCheng
Period 1	0.6638	0.7376	0.4958	0.6584	0.6592
Period 2	0.6199	0.6708	0.5053	0.6308	0.6426
Period 3	0.5388	0.5267	0.3503	0.7131	0.6549
Period 4	0.7000	0.5481	0.3807	0.6528	0.6269
S&P500	DAX	FTSE	NAS	ShangZheng	ShenCheng
Period 1	0.6510	0.7316	0.4818	0.6868	0.6604
Period 2	0.6033	0.6388	0.5378	0.6523	0.6553
Period 3	0.5472	0.5193	0.3041	0.7293	0.6649
Period 4	0.7181	0.5992	0.4127	0.6529	0.6336
ShangZheng	DAX	FTSE	NAS	S&P500	ShenCheng
Period 1	0.6587	0.7452	0.7041	0.6896	0.3370
Period 2	0.6121	0.6948	0.6982	0.6915	0.3990
Period 3	0.6863	0.6530	0.6823	0.7574	0.3572
Period 4	0.7217	0.6359	0.6697	0.6704	0.3500
ShenCheng	DAX	FTSE	NAS	S&P500	ShangZheng
Period 1	0.6533	0.7552	0.6953	0.7024	0.3550
Period 2	0.6344	0.6820	0.7327	0.7151	0.4150
Period 3	0.6463	0.6336	0.6485	0.7540	0.4085
Period 4	0.7123	0.6423	0.6379	0.6511	0.3494

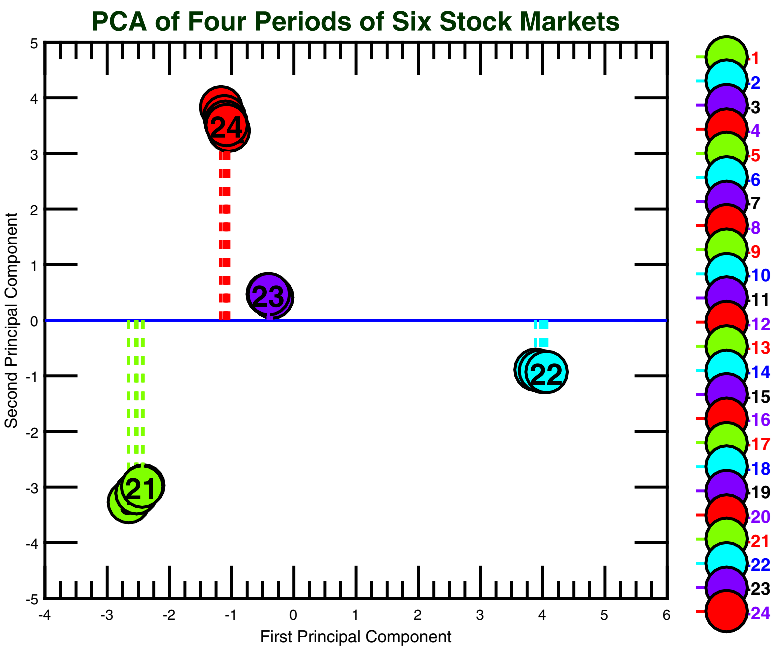

Fig. 7. The number labels 1,2,3,4 respectively correspond to four periods of the stock DAX, the number labels 5,6,7,8 respectively correspond to four periods of the stock FTSE, the number labels 9,10,11,12 respectively correspond to four periods of the stock NAS, the number labels 13,14,15,16 respectively correspond to four periods of the stock S&P500, the number labels 17,18,19,20 respectively correspond to four periods of the stock ShangZheng, the number labels 21,22,23,24 respectively correspond to four periods of the stock ShenCheng.

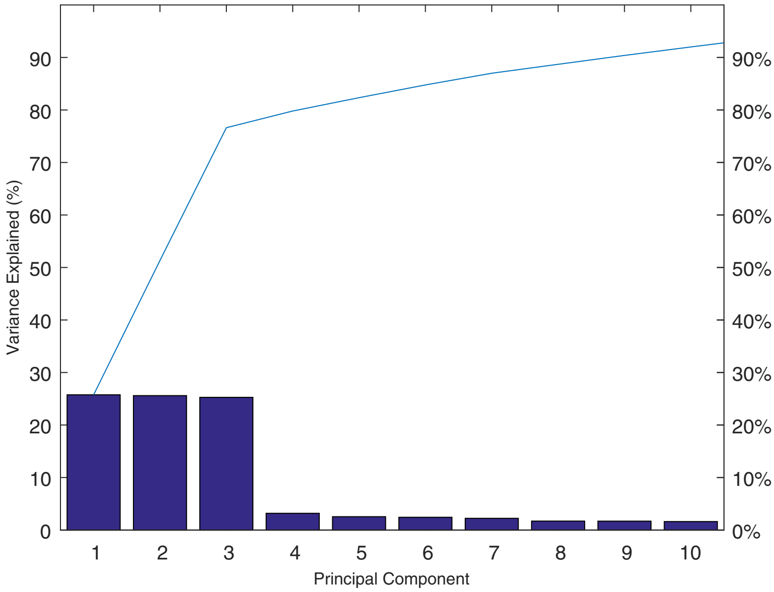

Fig. 8. The variance explained of Principal Component.

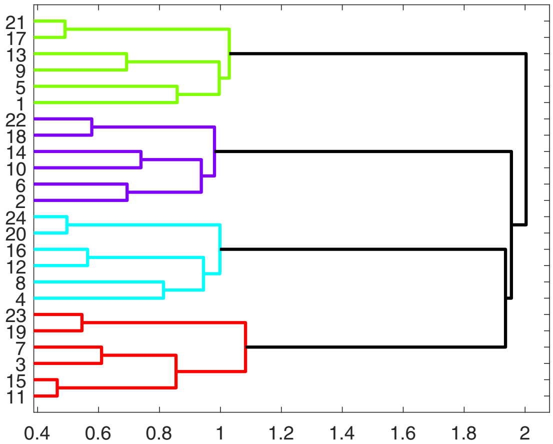

Fig. 9. Phylogenetic tree generated according to SPTE of stock indices during four periods.

4. Analysis and results

4.1. Dependence of SPTE on data size

First, in order to test the dependence of SPTEs and STEs reliability on the length of the time series, we study SPTE and STE between two signals by changing the lengths of data points. The SPTE and STE results of artificial signal generated by AR model, ARFIMA model and Henon map are represented in Fig. 5. The SPTE and STE values between two generated series are positively related to the data size. As desired, in both cases, as the data size increased, the information flow between the corresponding signals measured by SPTE and STE increased too. Regarding the AR model, ARFIMA model and Henon Map, as the data size increased, the increasing of entropy values obtained using SPTE are all lower than those obtained using STE (see Fig. 5). In other words, the proposed SPTE method is less sensitive to the data size and indicates short error bars. It can be seen that the information flow from X to Y is bigger than that from Y to X. This result is consistent with the findings of Eq. (9). Even when the data size is very large, the strength of the information flow is still not steady(see Fig. 5 (middle) and (right)).

In the following, we focus on the effect of noise on the SPTE and STE. We add different levels of noise to the two generated signals, and observe the changes of SPTE (see Fig. 6). In order to prevent the results only for a certain length of sequence is established, we select the 5 groups with different length sequences as contrast. In the previous part, we observe that SPTE curve tends to be stable when the sequence length is more than 30000, so our choice for the length of data is

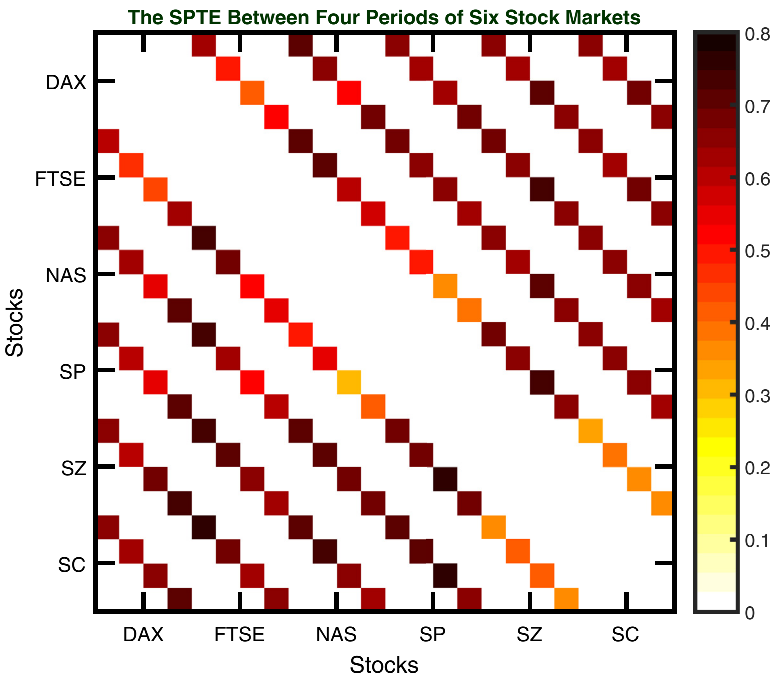

Fig. 10. The SPTE Between Four Periods of Six Stock Markets.

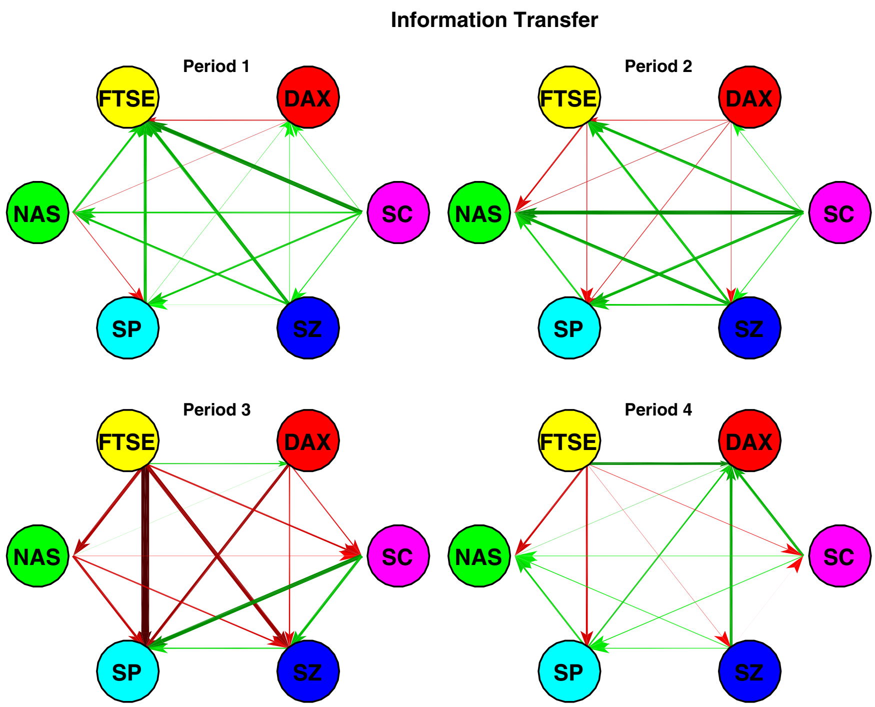

Fig. 11. Network representation of information flow. Network nodes represent stock markets and network links indicate the degree of interactions for pairs of stock markets. We calculate the information flow direction by $dSPTE(dSPTE_{(X \to Y)} = SPTE_{(X \to Y)} - SPTE_{(Y \to X)})$. If $dSPTE_{(X \to Y)}$ is positive, the line color is red, otherwise is green. The line thickness and darkness corresponds linearly to link strength. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

suitable. We can address from the following figures that the SPTE values with different levels of SNR noise in each case do not change significantly. Specifically, for the AR model, the influence of noise for the large amount of data and small amount of data sequence is relatively large. For the ARFIMA model and the Henon map, the change of the SNR has a great influence when data size is small.

4.2. The symbolic phase transfer entropy between the real financial data

To further characterize the direction and strength of information flow among stocks, we choose four significant periods for analysis. Then, for each period we respectively calculate SPTE (see Table 2) between each pair of stocks. For example, the first table in Table 2 is the values of SPTE from DAX to the other stocks. Probing the information flow among stocks, we discover that the SPTE between the two stocks from the same country is far less than the SPTE between stocks from different areas. The two stocks from the same country are governed by the same mechanism, while the stocks from different countries are based on different mechanisms. Accordingly, there is a stronger flow of information between different countries than that between the same country. The following analysis also confirms this conclusion. First, we use PCA method to verify the effectiveness of the proposed method. Then, we draw the Phylogenetic tree to address the similarity among six stocks during four different periods. Then, we show the direction and strength of information flow among six stocks during four periods in different ways.

4.3. Principal component approach(PCA) analysis

First, for further testing we use the principal component approach to verify these results [29]. Principal component analysis (PCA) uses an orthogonal transformation to transform a group of observations that are possibly correlated variables into a set of values that are linearly uncorrelated variables called principal components, and it is a statistical procedure. The results of PCA are shown in Figs. 7 and 8. Obviously, four periods of six stock markets are distinguished, and the clustering results by using PCA method suggests that SPTE could be used for reliable computation of the direction and strength of the information flow. In Fig. 9, we show the result of a Phylogenetic tree for the similarity among stock time series [30]. We find that six stock indices of four periods drop into four clusters, corresponding to four periods selected before. Furthermore, no matter during which period, two stocks of the same country are on the same branch. Those are in conformity with the practical situation. It suggests that four stocks from America and Europe have relatively high degree of similarity.

4.4. The symbolic phase transfer entropy analysis

For the sake of getting more effective information from Table 2, we present the more direct-viewing comparison of the values of SPTE in Fig. 10. Since we compare values of SPTE among the six stocks for the same period, there are a lot of map areas that are white, that is, the value of SPTE is 0. Fig. 10 displays the information flows of 24 subsections of stock markets. Note that for all the sub-periods, the SPTE values between the same country are relatively smaller than the SPTE of different countries. These results are consistent with the actual experience.

However, for the other part, it is hard for providing clear information. To probe the interaction behaviors in details, we decide to construct the network representation in Fig. 11. In 2008, the financial crisis broke out in the United States, but the crisis is in fact, from the beginning of the spring of 2007. The whole process of the crisis is period 1 and period 2 we choose before. Fig. 11 reveals that in those two periods, the two stocks of China have a great impact on the other stocks, which is also related to the good macroeconomic regulation and control of the government of China. In other words, the Chinese market may be playing a positive effect. In period 3, Europe has significant impact on other countries. Since 2009, the sovereign debt crisis broke out in some European countries. The European stock market has a great impact on other countries. Such as China, the first is the impact on exports. Second, the outbreak of European sovereign debt exacerbated the volatility of short-term capital flows. Besides, the influence of the two stocks of China for S&P500 is really dominant. Unlike the previous three periods, the period 4 is relatively stable. The difference of the impact between the stock is not very obvious during period 4. It is interesting that the impact of other stocks on DAX is relatively bigger. That result may partly originated in the fact that DAX market can be easily influenced by the other stocks.

5. Conclusion

In this paper, we propose the symbolic phase transfer entropy method to quantify the strength and direction of information flow. In order to test the capability of the proposed method, we study the SPTE of real stock markets and three artificial time series(AR model, ARFIMA model and Henon Map). In the experiment of the artificial data, we find that SPTE is less sensitive to data size and noise. The relative changing rates of SPTE are smaller than that of STE when data sizes are increasing. And the changes of SPTE values with different levels of SNR noise are not obvious. In the analysis of real-world stock markets, we calculate the symbolic phase transfer entropy for the daily closing prices of DAX, FTSE, NAS, S&P500, ShangZheng, ShenCheng over four periods. We test the effectiveness of SPTE by PCA techniques. We draw a rooted tree according to the result of the symbolic phase transfer. The six stock indices of four periods drop into four clusters, which in accordance with the four periods we selected before. And the results suggest that the SPTE is an effective method to quantify information flow among systems. In order to facilitate the analysis, we represent the symbolic phase transfer entropy by network diagram. We are able to observe that the transfer of information between two stocks from the same country is much smaller than the transfer of information from different countries. Information transfer between the stocks from different countries may vary in four periods. The two Chinese stock markets indicate strong influence on other countries when the financial crisis happened in 2008. When the European sovereign debt crisis broke, the impact of two European stocks on other stocks is relatively strong. Depending on the specific situation, this effect may be positive or negative. However, further studies will be needed to find out regulation mechanism of stock markets. It may be also of interest in the future to predict the stock according to the information flow.

Acknowledgements

We acknowledge financial support from the National Natural Science Foundation of China (61673005, 61304145 and 61371130).

References

[1] Zhao X, Shang P, Wang J. Measuring information interactions on the ordinal pattern of stock time series. Phys Rev E 2013;87(2):022805.
[2] Ahlswede R, Cai N, Li S-YR, Yeung RW. Network information flow. Inf Theory IEEE Trans 2000;46(4):1204–16.
[3] Fax JA, Murray RM. Information flow and cooperative control of vehicle formations. Autom Control, IEEE Trans 2004;49(9):1465–76.
[4] Lin A, Shang P, Zhong B. Hidden cross-correlation patterns in stock markets based on permutation cross-sample entropy and pca. Physica A 2014;416:259–72.
[5] Tian Q, Shang P, Feng G. Financial time series analysis based on information categorization method. Physica A 2014;416:183–91.
[6] Kwon O, Yang J-S. Information flow between composite stock index and individual stocks. Physica A 2008;387(12):2851–6.
[7] Schreiber T. Measuring information transfer. Phys Rev Lett 2000;85(2):461.
[8] Kaiser A, Schreiber T. Information transfer in continuous processes. Physica D 2002;166(1):43–62.
[9] Verdes P. Assessing causality from multivariate time series. Phys Rev E 2005;72(2):026222.
[10] Lungarella M, Pitti A, Kuniyoshi Y. Information transfer at multiple scales. Phys Rev E 2007;76(5):056117.
[11] Vakorin VA, Kovacevic N, McIntosh AR. Exploring transient transfer entropy based on a group-wise ica decomposition of eeg data. Neuroimage 2010;49(2):1593–600.
[12] Barnett L, Barrett AB, Seth AK. Granger causality and transfer entropy are equivalent for Gaussian variables. Phys Rev Lett 2009;103(23):238701.
[13] Sandoval L. Structure of a global network of financial companies based on transfer entropy. Entropy 2014;16(8):4443–82.
[14] Vicente R, Wibral M, Lindner M, Pipa G. Transfer entropyla model-free measure of effective connectivity for the neurosciences. J Comput Neurosci 2011;30(1):45–67.
[15] Nichols J, Seaver M, Trickey S, Todd M, Olson C, Overbey L. Detecting nonlinearity in structural systems using the transfer entropy. Phys Rev E 2005;72(4):046217.
[16] Wibral M, Rahm B, Rieder M, Lindner M, Vicente R, Kaiser J. Transfer entropy in magnetoencephalographic data: quantifying information flow in cortical and cerebellar networks. Prog Biophys Mol Biol 2011;105(1):80–97.
[17] Staniek M, Lehnertz K. Symbolic transfer entropy. Phys Rev Lett 2008;100(15):158101.
[18] Lobier M, Siebenhühner F, Palva S, Palva JM. Phase transfer entropy: a novel phase-based measure for directed connectivity in networks coupled by oscillatory interactions. Neuroimage 2014;85:853–72.
[19] Bandt C, Pompe B. Permutation entropy: a natural complexity measure for time series. Phys Rev Lett 2002;88(17):174102.
[20] Le Van Quyen M, Foucher J, Lachaux J-P, Rodriguez E, Lutz A, Martinerie J, et al. Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony. J Neurosci Methods 2001;111(2):83–98.
[21] Zhu J, Le Bouquin Jeannès R, Yang C, Bellanger J-J, Shu H. Partial mutual information for simple model order determination in multivariate eeg signals and its application to transfer entropy. In: Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on. IEEE; 2012. p. 673–6.
[22] Podobnik B, Grosse I, Horvatić D, Ilic S, Ivanov PC, Stanley HE. Quantifying cross-correlations using local and global detrending approaches. Eur Phys J B 2009;71(2):243–50.
[23] Granger CW. Long memory relationships and the aggregation of dynamic models. J Econom 1980;14(2):227–38.
[24] Granger CW, Joyeux R. An introduction to long-memory time series models and fractional differencing. J Time Series Anal 1980;1(1):15–29.
[25] Hosking JR. Fractional differencing. Biometrika 1981;68(1):165–76.
[26] Podobnik B, Ivanov PC, Jazbinsek V, Trontelj Z, Stanley HE, Grosse I. Power-law correlated processes with asymmetric distributions. Phys Rev E 2005;71(2):025104.
[27] Podobnik B, Ivanov PC, Biljakovic K, Horvatic D, Stanley HE, Grosse I. Fractionally integrated process with power-law correlations in variables and magnitudes. Phys Rev E 2005;72(2):026121.
[28] Wang X, Shang P, Huang J, Feng G. Data discretization for the transfer entropy in financial market. Fluctuation Noise Lett 2013;12(04):1350019.
[29] Jolliffe I. Principal component analysis. Wiley Online Library; 2002.
[30] Yin Y, Shang P. Modified multidimensional scaling approach to analyze financial markets. Chaos 2014;24(2):022102.