Commun Nonlinear Sci Numer Simulat 47 (2017) 338–347

Contents lists available at ScienceDirect

Commun Nonlinear Sci Numer Simulat

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

Research paper

# Transfer entropy between multivariate time series

Xuegeng Mao, Pengjian Shang*

*Department of Mathematics, School of Science, Beijing Jiaotong University, Beijing 100044, PR China*

ARTICLE INFO

*Article history:*
Received 27 April 2016
Revised 5 October 2016
Accepted 2 December 2016
Available online 5 December 2016

*Keywords:*
Multivariate time series
Transfer entropy
Time-delay reconstruction of phase space

ABSTRACT

It is a crucial topic to identify the direction and strength of the interdependence between time series in multivariate systems. In this paper, we propose the method of transfer entropy based on the theory of time-delay reconstruction of a phase space, which is a model-free approach to detect causalities in multivariate time series. This method overcomes the limitation that original transfer entropy only can capture which system drives the transition probabilities of another in scalar time series. Using artificial time series, we show that the driving character is obviously reflected with the increase of the coupling strength between two signals and confirm the effectiveness of the method with noise added. Furthermore, we utilize it to real-world data, namely financial time series, in order to characterize the information flow among different stocks.



## 1. Introduction

One of the biggest problems in scientific studies is modeling and analysis of the cause-and-effect relationships among dynamical system components. The research of interdependence among different observations addresses three major tasks: the formulation form of coupling, the direction and the quantification of coupling strength [1]. Over the last few years, researchers have introduced a lot of linear and non-linear analysis approach to detect coupling from response observations. Among them, Wiener proposed the leading approach to identify if the prediction of a given series is improved by the incorporation of information from the past of another series, which says: "For two simultaneously measured signals, if we can predict the first signal better by using the past information from the second one than by using the information without it, then we call the second signal causal to the first one." [2]. And later Granger formalized it, called Granger causality [3]. Granger causality is robust in the study of quantifying the directional interactions and causal relationships among time series in dynamical systems and it has been universally applied in economics. However, it is based on linear models and is not appropriate for the nonlinear system. Lots of extensions of Granger's idea have thus been proposed, such as conditional Granger causality [4] and multivariate Granger causality [5]. In spite of the success of these strategies, the model-based approach may suffer from the drawbacks of model misspecification and result in unauthentic information of causality.

Recently, Podobnik and Stanley introduced a new measurement termed as detrended cross-correlation analysis(DCCA) to quantify the power low cross-correlation between non-stationary time series [6,7], which is a extension of detrended fluctuation analysis(DFA) [8] and is based on detrended covariance. Based on the theory of DFA and DCCA, the detrended cross-correlation coefficient $\rho_{DCCA}$ was proposed to quantify the level of cross-correlation between non-stationary signals [9]. Furthermore, the generalization of $\rho_{DCCA}$, q-dependent detrended cross-correlation coefficient $\rho_q$, was introduced by Jarosław Kwapień [7], in which way the insensitivity of$\rho_{DCCA}$could be avoid. Nowadays these methods have become a

* Corresponding author. Tel.: +8613641228092
*E-mail addresses:* 12271067@bjtu.edu.cn (X. Mao), pjshang@bjtu.edu.cn (P. Shang).

http://dx.doi.org/10.1016/j.cnsns.2016.12.008
1007-5704/

widely-used tool for investigating the cross-relations between time series, spanning from finance, physiology, astrophysis and so on [10,11]. However, the method it doesn't represent the directionality between different time series, too.

In the content of information theory, the concept of entropy was first put forward by Shannon [12], which aims to decrease the uncertainty of information. And several techniques have been utilized to identify the relationship between two random variables. Mutual information (MI) is one of the model-free methods [13]. MI tells us that how much information one random variable obtains from another. Not dependent on the real-value random variables like the distribution type of variables, MI is more natural and can estimate the similarity between the joint distribution and the products of factored marginal distribution. Nevertheless, the causal interactions cannot be identified by calculating MI due to lacking the definition of direction and dynamical properties. Because there is a major shortcoming of MI: MI is symmetric, that means, whatever X is a deterministic function of Y or Y is the driving character of X, there is no difference of information transition. To obtain more appreciate mechanism, delayed mutual information is taken into consideration by introducing a time-lag in one of the variable. Although delayed MI is an asymmetric measure, it consists in the certain dynamical framework as a result of the shared history. To this end, transfer entropy(TE) as the implementation of Wiener's theory was expressed by Schreiber [14].

Transfer entropy has been proved as an effective scalar measure of capturing the directional and dynamical features among different components of time series. This approach was defined originally for data by calculating discrete probability [14], and later added the algorithm of continuous variables [15]. Compared to model-based Granger causality and mutual information, TE is practical for both linear and non-linear systems and needn't presume any specific model for the causal effects of two systems. What's more, it is asymmetric and built upon transition probability. Since its mechanism, TE has been applied to a variety of scientific fields, including economic [16,17], biological [18,19], chemical analysis [20], health detecting [21] and music [22]. In the last few years, more and more papers have concentrated on the interactions of neurosciences by means of TE [23].

Although TE is a rigorous measure within the information theoretic construction, it is limited to the scalar time series in the present papers. And some of them have just identified the causal inter-dependencies among three or more observations in a multivariate system [24,25]. What if the variables have two or more dimensions? After all multivariate time series are not rare in practical cases, for instance, economic data and biological data are usually high-dimensional [26]. But how could we deal with the variables so that we can calculate the transfer entropy between multidimensional variables? There are few studies thinking about a concept of coupling relationship in the multivariate system. Nevertheless, The technique of time-delay reconstruction of a phase space opens a new avenue for understanding the causal relationships between multivariate time series [27,28]. This measure has been utilized to the prediction of nonlinear systems and modeling [29,30], such as noise reduction [31], signal categorization [32] and control [33]. At the beginning, most of the published articles emphasize on scalar time series with the benefit of embedding theorem, due to the reason that scalar time series are enough for the reconstruction of dynamical systems if the delayed parameters are suitable. But it might be diverse in practical cases. If there are three or more dimensions in some system, we cannot use the third coordinate to reconstruct the dynamical system by reason that it does not resolve the former two coordinates' symmetry. Furthermore, there are significant merits using different time series to predict, extremely when the system is noisy. Now we can apply this technique to estimate the interactions among multivariate time series of system combined with the theory of transfer entropy.

The algorithm of calculating multivariate transfer entropy combines the theory of time-delay reconstruction of a phase space and the traditional measurement of transfer entropy. Considering that most of the practical time series are contaminated with a noise, it can be widely used in natural complex system. For example, there are numerous parameters in traffic system, such as traffic speed, volume and occupancy [34]. They together represent the current condition of a certain road. To analyze the inter-dependency among different roads in the round, it is better to adopt two or more parameters. What's more, in stock markets, it is obvious that there is an inextricable connection between the closing price and volume on the trading day. We can obtain more comprehensive conclusions about causality among different stock indices by using both.

The structure of the paper is as follows. Section 2 first illustrates the concept of phase space reconstruction and transfer entropy, then presents the algorithm of multivariate transfer entropy based on the theory of phase space reconstruction. In Section 3 we use two types of artificial time series, the unidirectionally coupled Rössler system and Hénon map, to examine the effectiveness of the fresh technique, also including the effects of different types noise added in the data. Section 4 demonstrates the and application of financial time series. Finally, Section 5 summarizes the conclusions.

## 2. Methodologies

### 2.1. Time-delay reconstruction of a phase space

Consider an M-dimensional time series $X=(X_1, X_2,\ldots, X_N)^T$, where $X_i=(x_{i,\,1}, x_{i,\,2},\ldots, x_{i,\,M}), i=1, 2,\ldots, N$, then using time-delay embedding theory [27,28] as in the case of scalar time series, we make a time-delay reconstruction:

$$\begin{aligned} W_k = \big(&x_{1,k}, x_{1,k+\tau_1}, \ldots, x_{1,k+(d_1-1)\tau_1},\\ &x_{2,k}, x_{2,k+\tau_2}, \ldots, x_{2,k+(d_2-1)\tau_2},\\ &\quad\cdots \qquad \cdots\\ &x_{N,k}, x_{N,k+\tau_M}, \ldots, x_{N,k+(d_N-1)\tau_N}\big), \end{aligned} \tag{1}$$

where $\tau=(\tau_1, \tau_2,..., \tau_N)$ and $D=(d_1, d_2,..., d_N)$ are time-delay and embedding dimension respectively. Next issue is that how we choose the acceptable time-delays $\tau_i$ and embedding dimensions $d_i$, $i=1, 2,..., N$. We can use mutual information [35] and auto-correlation [36] to choose the time-delay for a scalar time series. But the present question is how to calculate $\tau_i$ separately for every scalar time series. It is of great importance to determine a perfect time-delay because different time-delays may lead to magnificent affects for the time series. In this paper, we use mutual information to choose the time-delays. Having already chosen the time-delays, we can find the embedding dimensions by means of Cao algorithm or FNN. The concrete calculating process sees Ref [27].

### 2.2. Transfer entropy

Transfer entropy (TE) is a non-parametric measure that estimates the directed information flow among stochastic processes, which is developed by Schreiber [14]. It is relevant to the concept of Shannon entropy [12] and mutual information [13] in the framework of information theory. More importantly, transfer entropy is a powerful technique quantifying the coupling strength and asymmetric properties in dynamical systems.

Suppose a random system with sets of possible issues whose probabilities of occurrence are $p_1, p_2,..., p_n$, we define $\log 1/p(i)$ as the information quantity of any certain event. Averaging this amount over all possible outcomes, we obtain the formulation

$$H = -\sum_i p_i \log p_i, \tag{2}$$

which is called the entropy. If there are two random variables in the system, the joint entropy girls

$$H(X,Y) = -\sum_{x\in X}\sum_{y\in Y} p(x,y)\log p(x,y) \tag{3}$$

is given to detect the properties of system. What's more, conditional entropy girls

$$H(X|Y) = -\sum_{y\in Y} p(y)\sum_{x\in X} p(x|y)\log p(x|y). \tag{4}$$

is to reduce the unpredictability of the other when two variables are inter-dependent with each other. Then mutual information(MI) is proposed to identify the inter-dependencies between them, which can be written as:

$$MI(X;Y) = -\sum_{x,y} p(x,y)\log p(x,y)/(p(x)p(y)). \tag{5}$$

Notice that

$$\begin{aligned} MI(X;Y) &= H(X) - H(X|Y) = H(X) + H(Y) - H(X,Y) \\ MI(Y;X) &= H(Y) - H(Y|X) = H(Y) - H(X) - H(X,Y) \end{aligned} \tag{6}$$

So $MI(X; Y)=MI(Y, X)$ (because $H(X, Y)=H(Y, X)$), that is why MI cannot provide the directional information even though it is a good measure to quantify the relationships between two observations. Lately, Schreiber proposed transfer entropy to characterize the coupling strength and directional properties from one system to the other.

Assume that $I$: $\{x_n, n=1, 2,..., N\}$ and $J$: $\{y_n, n=1, 2,..., N\}$ are two systems with $N$random variables. If the time series $I$ is a Markov process of degree $k$, we can obtain the equation:

$$p(x_{n+1}|x_n, x_{n-1}, \dots, x_1) = p(x_{n+1}|x_n, x_{n-1}, \dots, x_{n-k+1}), \tag{7}$$

which represents that the state $x_{n+1}$ is related to the $k$ previous states of itself. In a similar way, the state $y_{n+1}$ depends on the $l$ previous states in the system of $J$. To be specific, define $x_n^{(k)} = (x_n, x_{n-1}, \dots, x_{n-k+1})$ and $y_n^{(l)} = (y_n, y_{n-1}, \dots, y_{n-l+1})$ with the length of $k$ and $l$ respectively.

Considering the information transfer from $J$ to $I$, it is advisable to quantify the deviation from the generalized Markov property,

$$p(x_{n+1}|x_n^{(k)}) = p(x_{n+1}|x_n^{(k)}, y_n^{(l)}). \tag{8}$$

It is to say that the state of $J$ is independent on the transition probabilities on system $I$. Assume that the next state $x_{n+1}$ relays on both the $k$ previous states of the same variable and the $l$ previous states of system $J$, the formulation of transfer entropy is defined as:

$$\begin{aligned} T_{J\to I}(k,l) &= \sum p\left(x_{n+1}, x_n^{(k)}, y_n^{(l)}\right)\log\frac{p\left(x_{n+1}|x_n^{(k)}, y_n^{(l)}\right)}{p\left(x_{n+1}|x_n^{(k)}\right)} \\ T_{I\to J}(k,l) &= \sum p\left(y_{n+1}, x_n^{(k)}, y_n^{(l)}\right)\log\frac{p\left(y_{n+1}|x_n^{(k)}, y_n^{(l)}\right)}{p\left(y_{n+1}|y_n^{(k)}\right)} \end{aligned} \tag{9}$$

The joint probability density function $p(x_{n+1}, x_n^{(k)}, y_n^{(l)})$is the probability of combination of $x_{n+1}, x_n^{(k)}$ and $y_n^{(l)}$ with particular values. The conditional PDF $p(x_{n+1}|x_n^{(k)}, y_n^{(l)})$ and $p(x_{n+1}|x_n^{(k)})$ are the probability that $x_{n+1}$ has a certain value when the value of previous variables $x_n^{(k)}$ and $y_n^{(l)}$ are given and $x_n^{(k)}$ is given respectively. In the respect of practical cases, $k=1$ and $l=1$ are the better choices.

The transfer entropy from $J$ to $I$ quantifies the dynamical information of process $J$influencing the transition probabilities of another system $I$, and its asymmetric characteristic allows it to provide the information about the direction of inter-dependency between two time series.

### *2.3. Multivariate transfer entropy based on time-delay reconstruction of a phase space*

The innovative measure to estimate inter-dependencies between multi-dimensional variables of two system components is built upon the reconstruction of phase space [27], by which numbers of scalar time series are obtained. After that, the transfer entropy among time series can be calculated. This method consists of five steps. Consider two M-dimensional time series:

$$X = (X_1, X_2, \ldots, X_N)^T \text{ and } Y = (Y_1, Y_2, \ldots, Y_N)^T$$

where $X_i=(x_{i,\,1}, x_{i,\,2},\ldots, x_{i,\,M})$, $Y_i=(y_{i,\,1}, y_{i,\,2},\ldots, y_{i,\,M})$, $i=1, 2,\ldots, N$.

**Step 1:** Calculate the time-delay $\tau=(\tau_1, \tau_2,\ldots, \tau_N)$and embedding dimension $D=(d_1, d_2,\ldots, d_N)$of X and Y referencing to Section 2.1.

**Step 2:** Make a time-delay reconstruction

$$\begin{aligned} P_k = (&x_{1,k}, x_{1,k+\tau_1}, \ldots, x_{1,k+(d_1-1)\tau_1}, & Q_k = (&y_{1,k}, y_{1,k+\tau_1}, \ldots, y_{1,k+(d_1-1)\tau_1},\\ &x_{2,k}, x_{2,k+\tau_2}, \ldots, x_{2,k+(d_2-1)\tau_2}, & &y_{2,k}, y_{2,k+\tau_2}, \ldots, y_{2,k+(d_2-1)\tau_2},\\ &\ldots \quad \ldots & &\ldots \quad \ldots\\ &x_{N,k}, x_{N,k+\tau_M}, \ldots, x_{N,k+(d_N-1)\tau_N}) & &y_{N,k}, y_{N,k+\tau_N}, \ldots, y_{N,k+(d_N-1)\tau_N}) \end{aligned} \tag{10}$$

where $k \le M - \max(d_j) \times \max(\tau_j)$ *and* $k+(d_j-1)\tau_j \le M$, $j=1, 2,\ldots, N$.

**Step 3:** Calculate transfer entropy $TE_{Q_k\to P_k}$ *and* $TE_{P_k\to Q_k}$ between $P_k$ and $Q_k$ according to. Section 2.2

**Step 4:** In order to calculate the transfer entropy between X and Y, we average the several TEvalues of $TE_{Q_k\to P_k}$ *and* $TE_{P_k\to Q_k}$ obtained in step 3.

$$\begin{aligned} TE_{Y\to X} &= \left(\sum_k TE_{Q_k\to P_k}\right)/k_{\max},\\ TE_{X\to Y} &= \left(\sum_k TE_{P_k\to Q_k}\right)/k_{\max}, \text{ where } k_{\max} = M - \max(d_j) \times \max(\tau_j) \end{aligned} \tag{11}$$

**Step 5:** Sometimes the results are small scale. To overcome the problem and summarize in avisualized way about the information of casual relationships, some researchers apply thenormalized directionality transfer entropy (NDTE) [37]. It is disposed by:

$$TE = (TE_{Y\to X} - TE_{X\to Y})/(TE_{Y\to X} + TE_{X\to Y}). \tag{12}$$

The value of NDTE varies from $-1$ to1. And the value is expected to be positive when $X$is the driver and negative for $Y$ driving $X$.

## 3. Numerical results for artificial time series

In this section, we apply two types of artificial time series, the unidirectionally coupled Rössler system [38,39] and two-way coupled Hénon map [40], to examine the effectiveness of the fresh technique and make some comparisons.

### *3.1. Unidirectionally coupled Rössler system*

The Rössler system is an oscillatory coupled system and is identified by

$$\begin{aligned} \dot{x}_1 &= -\omega_1 x_2 - x_3,\\ \dot{x}_2 &= \omega_1 x_1 + 0.15 x_2,\\ \dot{x}_3 &= 0.2 + x_3(x_1 - 10) \end{aligned} \tag{13}$$

for the driving system and

$$\begin{aligned} \dot{y}_1 &= -\omega_2 y_2 - y_3 + \varepsilon(x_1 - y_1),\\ \dot{y}_2 &= \omega_2 y_1 + 0.15 y_2,\\ \dot{y}_3 &= 0.2 + y_3(y_1 - 10) \end{aligned} \tag{14}$$

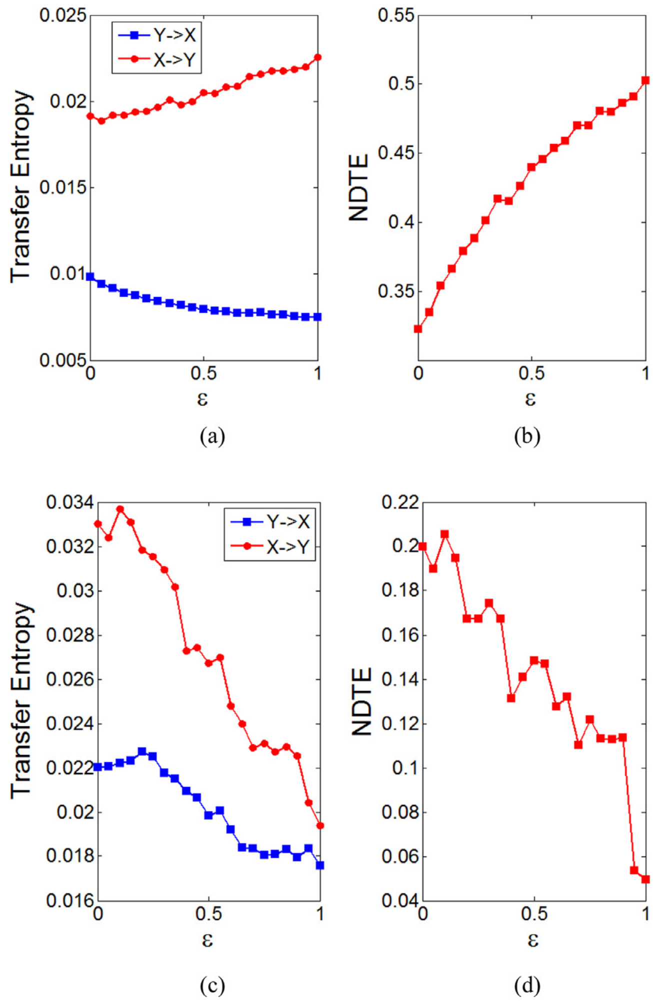


**Fig. 1.** (a) the transfer entropy from $X$ to $Y$ and from $Y$ to $X$ and (b) the normalized directionality transfer entropy (NDTE) between $X$ and $Y$ on different coupling strength $\varepsilon$ in the Rossler system with the frequencies $\omega_1 = 1.015$, $\omega_2 = 0.985$. (c) the transfer entropy from $X$ to $Y$ and from $Y$ to $X$ and (d) the normalized directionality transfer entropy (NDTE) between $X$ and $Y$ on different coupling strength $\varepsilon$ in the Rossler system with the frequencies $\omega_1 = 0.5$, $\omega_2 = 2.515$.

for the response system, where $\omega_1$ and $\omega_2$ are the mismatched parameters, $\varepsilon$denotes the strength of the coupling . The system was generated referring to Euler's method and afterwards interpolated at regular time intervals of 0.01 s. The duration of the time series was 100 s. We assume that $\omega_1 = 1.015$, $\omega_2 = 0.985$ and $\omega_1 = 0.5$, $\omega_2 = 2.515$.

Fig. 1 illustrates the influence of transfer entropy on the detectability of coupling strength in Rössler oscillators where the coupling strength $\varepsilon$ is presented in the $x$ components of system. Automatic controlled system has its own track, and the track of response system is controlled by equation itself and driving system at the same time. In our simulations, the frequencies are set as $\omega_1 = 1.015$, $\omega_2 = 0.985$ and $\omega_1 = 0.5$, $\omega_2 = 2.515$respectively. Fig. 1(a) shows the dependence between $X$ and $Y$ with the frequencies of $\omega_1 = 1.015$, $\omega_2 = 0.985$. We observe that the driver $X$ shows the growing influence on the response $Y$ with increasing $\varepsilon$ despite the low degree of identification. Fig. 1(b) shows the values of normalized directionality transfer entropy (NDTE) between $X$ and $Y$ with different coupling strength. Obviously, with the increase of coupling strength, the driving element shows a noticeable reflection of the system, which agrees with the equations. Fig. 1(c) and (d) shows the dependence between $X$ and $Y$ with the frequencies of $\omega_1 = 0.5$, $\omega_2 = 2.515$. With increasing coupling strength, $TE_{X\to Y}$ and $TE_{Y\to X}$ both decreased. Additionally, the NDTE also presents a decreasing trend. This trend can be explained by the cardiorespiratory interactions with the frequency ratio 1:5. And the autonomous system with $\omega_1 = 0.5$ is chaotic, while the autonomous system with $\omega_2 = 2.515$ is again quasiperiodic.

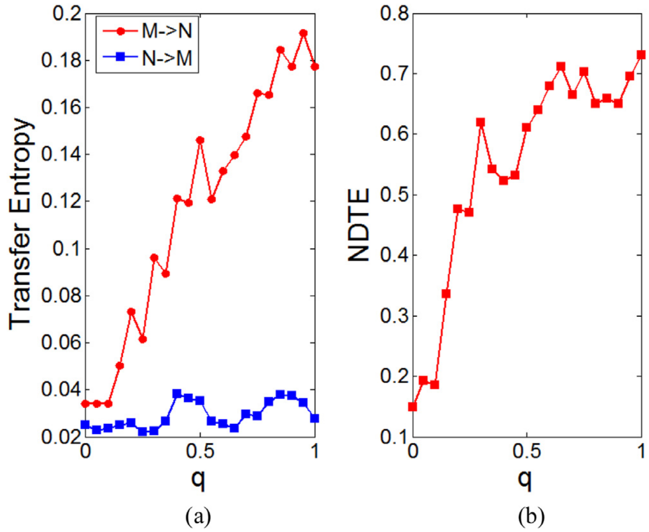


**Fig. 2.** (a) The transfer entropy from $M$ to $N$ and from $N$ to $M$ and (b) the normalized directionality transfer entropy (NDTE) between $M$ and $N$ on different coupling strength $q$ in Henon map.

### 3.2. Hénon map

The Hénon map is a two-dimensional dynamical system with discrete time series, which is proposed by Michel Hénon [40,41]. It is a simplification of Lorenz system with chaotic solutions. It is given by:

$$\begin{cases} x_t = 1 + y_{t-1} - a x_{t-1}^2 \\ y_t = b x_{t-1} \end{cases}, \tag{15}$$

where $a=1.4$, $b=0.3$, $t=1, 2,\ldots, N$. $N$ is the length of generated time series.

Suppose that $L_t=(x_t, y_t)^T$, $M_0=(r_1, r_2)^T$, $N_0=(s_1, s_2)^T$ ,where $r_i, s_i \sim N(0, 1), i = 1, 2$ and construct the two-dimensional time series with coupling strength $q$ as follow:

$$\begin{aligned} M_t &= 0.9M_{t-1} + L_{t-1} + M_0 \\ N_t &= 0.2N_{t-1} + qM_{t-1} + L_{n-1} + N_0 \end{aligned} \tag{16}$$

In general, Fig. 2 shows that the driving character of time series is obviously shown up with the increase of the coupling strength though there is a small fluctuation. From the Fig. 2(b), there are two inflection points at $q=0.4$ and $q=0.8$, where is relevant to the properties of two-way coupled system.

### 3.3. The effectiveness test of multivariate transfer entropy

In order to verify the effectiveness of transfer entropy between multivariate time series, we add two types of synthetic noise signals, white noise and the noise whose distribution is uniform, to generated data in Section 3.2. To be simplified, we use the results of normalized directionality transfer entropy(NDTE) as reference and calculate the deviate rate between the values of normalized directionality transfer entropy and that with noise.

First, we consider the time series contaminated by Gauss white noise whose standard deviations are varying from $\sigma=0.025$ to $\sigma=0.15$ in steps of 0.025. Fifty independent noise samples are applied in each simulation and the length of each noise sample is 1000. All results presented are the means of 50 observations of time series with the given level of noise and coupling strength.

Fig. 3 shows the deviate rates of normalized directionality transfer entropy with different levels of white noise and coupling strength. Overall, the range of deviate rate is within −5% to 5%, which indicates the noise added does not change the results considerably, although Hénon map is a little sensitive to changes. What's more, there are two obvious fluctuations at the point of $q=0.4$ and $q=0.8$ and the phenomenon is consistent with the conclusion in Section 3.2.

Furthermore, we add the noise whose distribution is uniform to the time series with different levels of noise by the equation

$$y_i = x_i(1 + \alpha p) \tag{17}$$

where $y_i$ is the time series with noise presented, $x_i$ is the original time series, $\alpha$ is the level of noise varying from 0.01 to 0.06 in steps of 0.01 and $p$ is a random number from the range −1 to 1. We also use fifty independent noise samples in each simulation and use the average of those results as final results with the given level of noise and coupling strength.

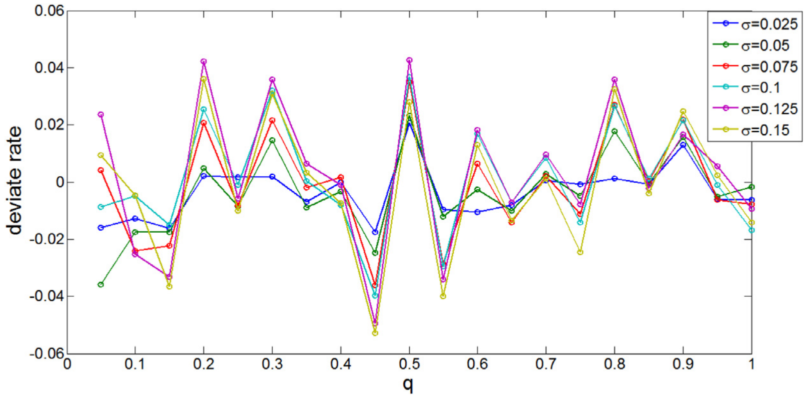


**Fig. 3.** The deviate rate of the normalized directionality transfer entropy (NDTE) between *M* and *N* with the white noise added.

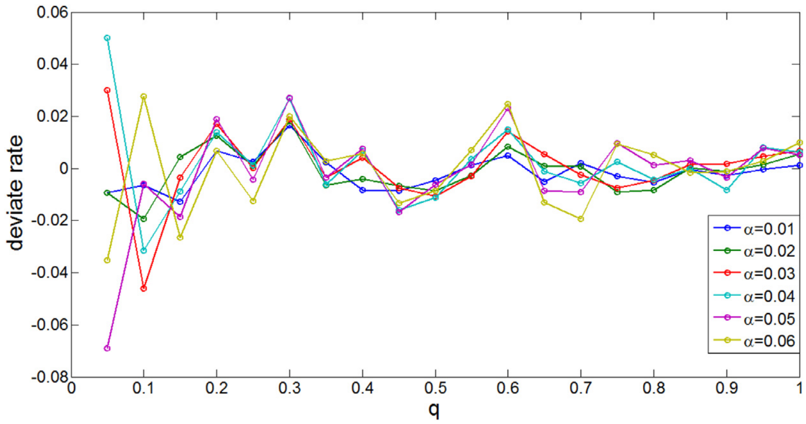


**Fig. 4.** The deviate rate of the normalized directionality transfer entropy (NDTE) between *M* and *N* with the noise whose distribution is uniform.

Fig. 4 illustrates the deviate rates of normalized directionality transfer entropy with different levels of noise whose distribution is uniform and coupling strength. We can see that the most values of deviate rates of normalized directionality transfer entropy are lower than those added white noise. As the increase of the level of noise, the deviate rate is close to 0 gradually despite of some fluctuations. In this case, the noise has little impact on the algorithm of calculating the transfer entropy between multivariate time series.

In general, adding different types of noise to time series has little impact on the results of transfer entropy between multivariate time series within a finite error range.

## 4. Numerical results for financial time series

To demonstrate the application of the introduced multivariate transfer entropy based on the reconstruction of phase space in practical situation, we utilize this measure to financial time series to characterize the information flow among them. Here, we consider the daily closing price and trading volume of 12 stock indices from January 1, 2011 to December 31, 2015. Every continent (America, Asia, Europe) contains 4 stock indices. The data used for this study was obtained from the website http://in.finance.yahoo.com/.

For the purpose of the synchronicity of time series, we remove the unwanted data and then recombine the time series. In this section, we adopt logarithmic price difference given by:

$$x_n \equiv \ln(S_n) - \ln(S_{n-1}), \tag{18}$$

where $S_n$ means the closing price of $n$th trading day.

As for trading volume $V_n$, we first standardize the data to keep the uniformity. Then we also use the logarithmic trading volume difference to deal with the standardized data in order to decrease the diversity between them.

In the experiment, our subject is to study the information flow among the 12 stock indices as follow.

Fig. 5 is the gray-scale map of the transfer entropy among stock indices with different variables. Numbers in x-axis denote the indices displayed in Table 1. The given direction of information flow is from x-axis to y-axis. Every indice is set to interact with the last 11 stock indices. The darker lattice represents the higher value of transfer entropy. Fig. 5(a) indicates

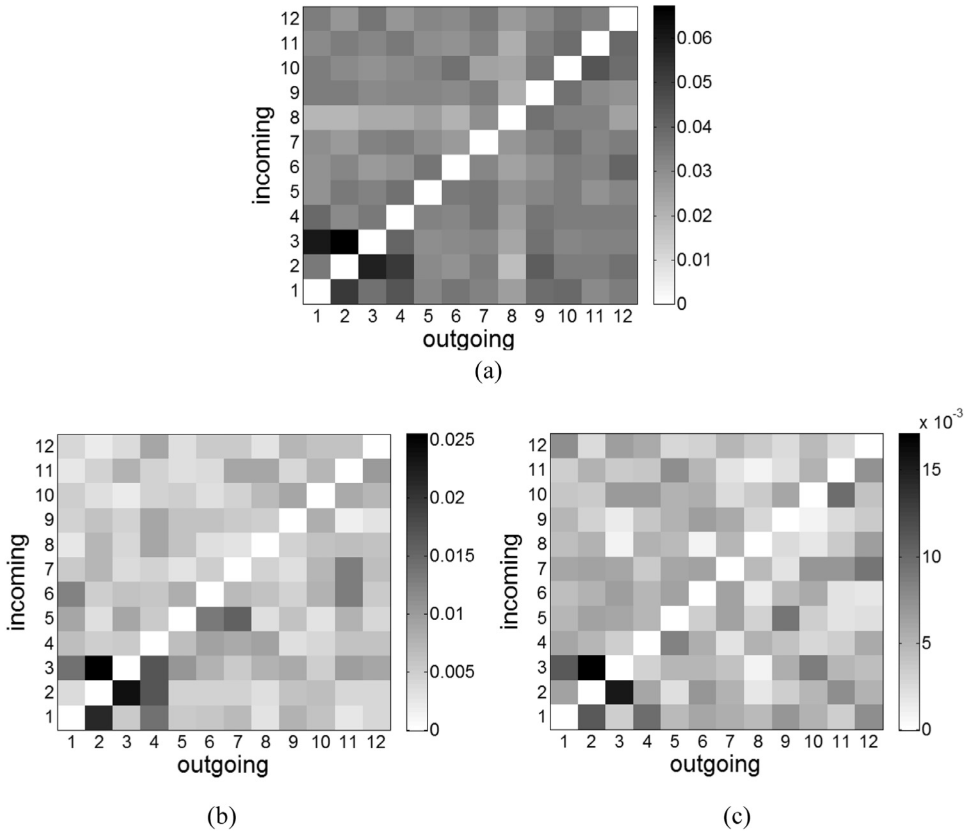


**Fig. 5.** (a) Transfer entropy for multivariate data, (b) transfer entropy for the closing price data, (c) transfer entropy for the trading volume data.

**Table 1**
List of 12 stock indices. We obtain data from the website http://in.finance.yahoo.com/.

| | | | |
|---|---|---|---|
| America | 1 | NDX | US |
| | 2 | NYA | US |
| | 3 | GSPC | US |
| | 4 | GSPTSE | Canada |
| Asia | 5 | HSI | China |
| | 6 | KS11 | Korea |
| | 7 | N225 | Japan |
| | 8 | BSESN | India |
| Europe | 9 | ATX | Austria |
| | 10 | FCHI | France |
| | 11 | FTSE | UK |
| | 12 | SSMI | Switzerland |

the transfer entropy by means of multivariate transfer entropy, that is to say, it takes the closing price and trading volume into consideration. Fig. 5(b) illustrates the transfer entropy of the closing price between stock indices. Fig. 5(c) shows the transfer entropy of the trading volume between stock indices. The stock markets in America drive most of the markets from other continents, especially the values of transfer entropy between markets in the US is extremely high, which is explained as the inner interactions and it is closely linked to high trading volume per day, while the transfer entropy for the market in Asia and Europe shows the lower values and demonstrably Num.8 (BSESN) has lower coupling inter-dependency between almost all the last listed market as a whole, in whose market the trading volume is also the lowest one.

To detect the efficiency of the proposed technique, we calculate the transfer entropy of the closing price and trading volume respectively, see Fig. 5(b) and (c). Compared Fig. 5(b) and (c) with Fig. 5(a), the two latter plots show little obvious coupling information among stock indices and the value gap of transfer entropy is a bit high. In summary, from Fig. 5, we can conclude that the markets in America occupy a leading position.

In some of the published dissertations, researchers have already studied the cross-correlations between any pair of stocks by means of DCCA and its extension. Results confirm that the cross-correlations are stronger if the pairs of stocks are in high level of industrial similarity [7]. Based on the conclusions above, we infer that the cross-correlations between stocks within the US are stronger than that in other continents daringly. However, the algorithm of detrended cross-correlation analysis between multivariate time series should be taken into consideration first.

## 5. Conclusion

In this paper, we initially introduce the method of multivariate transfer entropy based on the time-delay reconstruction of phase space when the variables of two systems are both two or high dimensional. The novelty is that this technique highlights the significance of constructing causality models able to quantify causal influence directly in non-stationary multivariate time series, which overcomes the limitation that the original transfer entropy only can capture which system drives the transition probabilities of another in scalar time series. In addition, the proposed method is relatively immune to additive noise.

To prove the robustness of this novel technique, the non-linear Rössler system and Hénon map first have been confirmed the effectiveness of the algorithm as the artificial time series. Moreover, we also confirmed the effectiveness of the proposed method by adding different types of noise to original time series. Then we applied it to financial time series to analyze the information flow between different stock markets. Here the innovation is that we combine the closing price and trading volume as the two-dimensional variables to estimate the causality relationships. By doing so, we overcame the inevitable problems when only analyzing the closing price and draw some more all-round conclusions.

Generally speaking, the proposed multivariate transfer entropy solved the problem concerning the information flow between multivariate time series successfully while there are a few shortcomings, e.g. one of the embedding dimensions is sometimes a small value, which may influence the calculation of transfer entropy.

## Acknowledgments

The financial supports from the funds of the China National Science (61371130) and the Beijing National Science (4162047) are gratefully acknowledged.

## References

[1] Papana A, Kyrtsou C, Kugiumtzis D, Diks C. Detecting causality in non-stationary time series using partial symbolic transfer entropy: evidence in financial data. Comput Econ 2015;47:341–65.
[2] Wiener N. The theory of prediction. Modern Math Eng 1956;1:125–39.
[3] Granger CW. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 1969:424–38.
[4] Geweke J. Measurement of linear dependence and feedback between multiple time series. J Am Stat Assoc 1982;77:304–13.
[5] Chen Y, Rangarajan G, Feng J, Ding M. Analyzing multiple nonlinear time series with extended Granger causality. Phys Lett A 2004;324:26–35.
[6] Podobnik B, Stanley HE. Detrended cross-correlation analysis: a new method for analyzing two non-stationary time series. Phys Rev Lett 2007;100:38–71.
[7] Kwapień J, Oświęcimka P, Drożdż S. Detrended fluctuation analysis made flexible to detect range of cross-correlated fluctuations. Phys Rev E Stat Nonlinear Soft Matter Phys 2015;92:052815.
[8] Peng CK, Buldyrev SV, Havlin S, Simons M, Stanley HE, Goldberger AL. Mosaic organization of DNA nucleotides. Phys Rev E Stat Phys Plasmas Fluids Related Interdisciplinary Topics 1994;49:1685–9.
[9] Zebende GF. DCCA cross-correlation coefficient: Quantifying level of cross-correlation. Physica A Stat Mech Appl 2011;390:614–18.
[10] Oświęcimka P, Drożdż S, Forczek M, Jadach S, Kwapień J. Detrended cross-correlation analysis consistently extended to multifractality. Phys Rev E Stat Nonlinear Soft Matter Phys 2014;89:023305.
[11] Rak R, Drozdz S, Kwapien J, Oswiecimka P. Detrended cross-correlations between returns, volatility, trading activity, and volume traded for the stock market companies. EPL (Europhys Lett) 2015;112:48001.
[12] Shannon CE. A mathematical theory of communication. ACM SIGMOBILE Mob Comput Commun Rev 2001;5:3–55.
[13] Gel′fand, I.M., ÎÂglom, A. Calculation of the amount of information about a random function contained in another such function, American Mathematical Society Providence; 1959.
[14] Schreiber T. Measuring information transfer. Phys Rev Lett 2000;85:461.
[15] Kaiser A, Schreiber T. Information transfer in continuous processes. Physica D 2002;166:43–62.
[16] Baek, S.K., Jung, W-S., Kwon, O., Moon, H-T. Transfer entropy analysis of the stock market, arXiv preprint physics/0509014. (2005).
[17] Marschinski R, Kantz H. Analysing the information flow between financial time series. Eur Phys J B Condensed Matter Complex Syst 2002;30:275–81.
[18] Rubinov M, Sporns O. Complex network measures of brain connectivity: uses and interpretations. Neuroimage. 2010;52:1059–69.
[19] Vicente R, Wibral M, Lindner M, Pipa G. Transfer entropy—a model-free measure of effective connectivity for the neurosciences. J Comput Neurosci 2011;30:45–67.
[20] Bauer M, Cox JW, Caveness MH, Downs JJ, Thornhill NF. Finding the direction of disturbance propagation in a chemical process using transfer entropy,. Control Syst Tech IEEE Trans 2007;15:12–21.
[21] 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;72:046217.
[22] Kulp CW, Schlingmann D. Using mathematica to compose music and analyze music with information theory. In: Mathematics and computation in music. Springer; 2007. p. 441–8.
[23] 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. Progr Biophys Mole Biol 2011;105:80–97.
[24] Runge J, Heitzig J, Marwan N, Kurths J. Quantifying causal coupling strength: a lag-specific measure for multivariate time series related to transfer entropy. Phys Rev E 2012;86:061121.
[25] Runge J, Heitzig J, Petoukhov V, Kurths J. Escaping the curse of dimensionality in estimating multivariate transfer entropy. Phys Rev Lett 2012;108:258701.
[26] Xiong H, Shang P. Weighted multifractal cross-correlation analysis based on Shannon entropy. Commun Nonlinear Sci Numer Simul 2016;30:268–83.
[27] Cao L, Mees A, Judd K. Dynamics from multivariate time series. Physica D 1998;121:75–88.
[28] Takens F. Detecting strange attractors in turbulence. Springer; 1981.

[29] Casdagli M. Nonlinear prediction of chaotic time series. Physica D 1989;35:335–56.
[30] Farmer JD, Sidorowich JJ. Predicting chaotic time series. Phys Rev Lett 1987;59:845.
[31] Grassberger P, Hegger R, Kantz H, Schaffrath C, Schreiber T. On noise reduction methods for chaotic data. Chaos 1993;3:127–41.
[32] Kadtke J. Classification of highly noisy signals using global dynamical models. Phys Lett A 1995;203:196–202.
[33] Ott E, Grebogi C, Yorke JA. Controlling chaos. Phys Rev Lett 1990;64:1196.
[34] Xu M, Shang P, Xia J. Traffic signals analysis using qSDiff and qHDiff with surrogate data. Commun Nonlinear Sci Numer Simul 2015;28:98–108.
[35] Fraser AM, Swinney HL. Independent coordinates for strange attractors from mutual information. Phys Rev A 1986;33:1134.
[36] Albano A-M, Mees A, De Guzman G, Rapp P. Data requirements for reliable estimation of correlation dimensions. Springer; 1987.
[37] Dimpfl T, Huergo L, Peter FJ. Using transfer entropy to measure information flows from and to the CDS market. In: Proc. of the European economic association and econometric society; 2011. p. 25–9.
[38] Liu Z. Measuring the degree of synchronization from time series data. EPL (Europhys Lett) 2004;68:19.
[39] Paluš M, Vejmelka M. Directionality of coupling from bivariate time series: how to avoid false causalities and missed connections. Phys Rev E 2007;75:056211.
[40] Hénon M. A two-dimensional mapping with a strange attractor. Commun Math Phys 1976;50:69–77.
[41] Xu M, Shang P, Huang J. Modified generalized sample entropy and surrogate data analysis for stock markets. Commun Nonlinear Sci Numer Simul 2016;35:17–24.