Article

Applying Permutation Transition Entropy to analyze the dynamical complexity of non-stationary time series

Xiaowei Mao [1], Chaoquan Jiang [2] and Zhigang Kou [2,*], Lizhang Chen [3]

[1] Affiliation 1; cqjiang@bjtu.edu.cn
[2] Affiliation 2; xwmao@bjtu.edu.cn
* Correspondence: cqjiang@bjtu.edu.cn; Tel.: +86-10-18800153987

Abstract: In this paper, we apply the method of Permutation Transition Entropy, which quantifies the Markov states transition between adjacent permutations, to measure the dynamical complexity of non-stationary time series. This method can capture the change of states trajectory of the underlying system by quantifying the Markov states transition between adjacent permutations. Unlike many traditional methods which usually measure the static complexity, the new method of permutation transition entropy (PTE) is able to identify the dynamical complexity with respect to the temporal structure change of the time series. By numerical analyses, we show that the PTE can give new information while other methods, like the permutation entropy (PE), cannot. We apply the PTE method to the financial time series, and find the existence of the momentum effect in the daily closing price and the daily trading volume of NASDAQ Composite Index. It indicates that the dynamical complexity of the index is lower than that of the purely random time series. The forthcoming state of the daily closing price is, therefore, a bit regular, which can be used for prediction. While the logarithm return shows very similar PTE values with those of the purely random time series, which correspond to high dynamical complexity. Furthermore, with multiscale analysis towards PTE, it turns out that under the same embedding dimension, the series with higher time-scale are more deterministic and represents more obvious momentum effect.

Keywords: Permutation Transition Entropy, Dynamical complexity, Markov state transition, Time series analysis, Multiscale analysis

1. Introduction

There have been many valued research about how to define the complexity of a complex system [1-3]. By measuring the irregularity of the output time series of the system [4,5], the complexity can be defined. When the system has irregular and diverse states, more information is required to reveal the transition of these states, which means the system has a high degree of static (structural) complexity.

The complexity of the system can also be measured in the dynamical scenario; it records the irregularly and diversity of the current state in the presence of past states [6-8]. If each system state cannot be easily predicted from the past states, so to describe the new system states requires a large amount of information, the system would have a high degree of complexity. However, if the system has few states, and the new system states can be easily predicted or described, it would have a low degree of complexity.

The dynamical complexity is closely related to the temporal structure of the time series. It refers to the intrinsic arrangement of the data with respect to the time. The temporal structure describes the connections between the neighboring data points and also distant data points in average. As for the linear connection, auto-correlation function is one of the simplest methods to estimate this kind of relationship. Analysis of the temporal structure can be used to characterize the complex system behaviors, to monitor the change of system states, and to predict the system trajectory. In physiology, the temporal structure of long-range correlated physiologic time series can robustly separate healthy and physiologic groups [9]. In finance, the temporal structure of stock time series is used to quantify the stock inefficiency and to distinguish the stage of stock market development [10,11]. To be clear, when the output time series of the system is regularly organized in the temporal structure, or at any time it is strongly correlated with its past values, then the series is more likely to be predicted [12,13]. It means that fewer information is needed for generating new data points, thus the time series has a low dynamical complexity.

There are many researches about detecting the complexity of a single time series. Also, there are many proposed measures, e.g.the Lyapunov exponent [14], fractal dimension [15], and entropy [16-20]. Most of the existing methods are classified as static complexity measures. It means that they measure the diversity of the system states, rather than the states change. Permutation entropy (PE) [21] is an information-theoretic complexity measure, which has the merits of the simplicity, extremely fast calculation, and robustness. But it also fails to characterize the dynamical complexity of the time series. Many researchers tried to detect the dynamical changes within time series by PE [22,23]. They used a moving window to estimate the local PE for a certain period, and observed the change of PE valued. But an additional parameter of the window size is required. In addition, this dynamical analysis cannot identify the instantaneous dynamical changes from one state to another one in time series as the window size should not be small [22].

In order to characterize the dynamical complexity of a single time series, we propose a new approach which capture the uncertainty of the permutation transition. The states of the underlying system can be described by the permutation motifs of the state vector. This paper uses a permutation transition entropy (PTE) to quantify the uncertainty of the permutations change. Unlike many traditional methods, such as the approximate entropy and sample entropy which compare the similarity or dissimilarity between long-distance state vectors [4], the PTE method recognizes the one-step states transition in the framework of the Markov chain. It is well-known as a stochastic model describing a sequence of possible events in which the probability of each event depends only on the state attained in the previous [24]. A low dynamical complexity is likely to have a small value of the PTE. Moreover, the PTE can reflect the predictability of the time series, since a lower dynamical complexity of the time series is likely to have a higher possibility of prediction.

We arrange the following paper as below. First, we retrospect the PE method, then introduce the PTE method. We discuss several analytic properties of the new method. Next, we test the PTE method with the artificial time series, then apply it to financial time series analysis, then we utilize multiscale analysis of PTE to further explore the dynamical complexity of single time series. Finally, we give our conclusions.

2. Materials and Methods

2.1. Permutation entropy(PE)

Consider the output time series of the underlying system:$\{x_t\}, t = 1, \ldots, T$. x_t is continuous, which means that equal values exist in a neighborhood can be neglected with no consequences because their probability of occurrence is rather small [25]. The deterministic(static) complexity of x_t can be estimated by the permutation entropy (PE) [21,26], which is a non-decreasing function of the entropy rate. The calculation of PE consists of the following steps.

1. At time t, we use the state vector X_t to describe the state of the system:

$$X_t = \{x_t, x_{t+\tau}, \ldots, x_{t+(m-1)\tau}\} \tag{1}$$

τ and m refer to the time delay and the embedding dimension in the phase space, respectively [27]. The matrix of state vectors at different time points is given by

$$X = \left(X_1, X_2, \ldots, X_{T-(m-1)\tau}\right)^{\Gamma} \tag{2}$$

Each state vector X_t has a unique permutation $\pi(X_t)$, by comparing the neighboring values in this vector. For example, the permutation of X_t= 3, 9, 4, 5 is $\pi(X_t) = 0231$, since $X_t(0) < X_t(2) < X_t(3) < X_t(1)$. At a given embedding dimension m, there exist at most m! permutations.

Count all types of permutations and number them π_i, respectively. Here i = 1,2,…, m!. For each $\pi(X_t) \in \{\pi_i | i = 1,2, \dots, m!\}$. The occurrence prob- ability of πi is approximated by the frequency:

$$P(\pi_i) = \frac{n\{X_t | X_t \ has \ permutation \ \pi_i\}}{T - (m-1)\tau} \tag{3}$$

Where $1 \leq t \leq T - (m-1)\tau$, and n represents the number of the set.

2. The information contained in π_i with the occurrence probability $p(\pi_i)$, can be quantified by $log2[1/p(\pi i)]$ in bit. The average amount of information of all the permutations is given by the expectation:

$$H_m = -\sum_{i=1}^{m!} P(\pi_i) log_2[P(\pi_i)] \tag{4}$$

H_m is known as the PE. PE quantifies the complexity and the irregularity of the data. H_m ranges between 0 and $\log_2 m!$). A large value of PE indicates a large degree of complexity of $\{x_t\}$. It attains to its maximum value when all the per- mutations are uniformly distributed, i.e. $P(\pi_i) = 1/m!$. It attains to its minimum value when one permutation dominates with the occurrence probability 1 while other permutations seldom arise. The PE can be normalized onto the range of [0,1], by:$h_m = H_m/(\log_2 m!)$ [28].

It is suggested by Bandt and Pompe [21] that the embedding dimension m is set to 3,4,5,6,7. It ensures that $m! \ll T$ satisfies. The time delay τ is usually set to 1. The advantages of the PE are its simplicity extremely fast calculation, robustness, and invariance with respect to nonlinear monotonous transformations.

2.2. Permutation transition entropy (PTE)

PE measures the static probabilities of the permutations. But the connection relationship between the neighboring permutations is not included. Here we incorporate the temporal structure of the time series, and study the dynamical transition probabilities of the permutations. A new dynamical complexity measure of permutation transition entropy (PTE) is proposed.

For the time series $\{x_t\}, t = 1, \dots, T,$ the state vector at time t is given by X_t, which has a unique permutation $\pi(X_t)$, where $\pi(X_t) \in \{\pi_i | i = 1, \dots, m!\}$. We estimate the dynamical complexity by quantifying the change of neighboring permutations. It is a one-step Markov state transition. Given a permutation motif, if the following permutation is more deterministic, the system would evolve in a more deterministic way. The system thus shows a low degree of complexity.

Suppose that $\pi(X_t) = \pi_i$, where $t = 1, \dots, T - (m-1)\tau$ and $i, j = 1, \dots, m!$ Here π_i can be equal to π_j. The one-step Markov transition probability from permutation π_i to its nearest neighbor π_j, is given by:

$$P(\pi_i \to \pi_j) = P(\pi_j | \pi_i) = \frac{p(\pi_i, \pi_j)}{p(\pi_i)} \tag{5}$$

The conditional probability $P(\pi_i, \pi_j)$ represents the transition probability of neighboring permutations. If one permutation is more likely to be followed by another permutation, that is, two permutations are likely to join together, the system would be more deterministic. It refers to a low degree of the dynamical complexity, and a high possibility of prediction. On the contrary, if one permutation is followed by different permutations at different time, the system would be more stochastic. It then refers to a high degree of dynamical complexity, and a low possibility of prediction.

PTE is defined as the conditional entropy, by summarizing all the permutations change:

$$T_m = -\sum_{i=1}^{m!}\sum_{j=1}^{m!} P(\pi_i, \pi_j) log_2[P(\pi_j|\pi_i)]$$

$$= -\sum_{i=1}^{m!}\sum_{j=1}^{m!} P(\pi_i, \pi_j) log_2 \frac{P(\pi_i, \pi_i)}{P(\pi_i)} = H_m(\pi_i, \pi_j) - H_m(\pi_i) \quad (6)$$

T_m is the PTE, $H_m(\pi_i, \pi_j)$ is the joint permutation entropy, and $H_m(\pi_i)$ is the PE. PTE quantifies the amount of information needed to describe permutation π_j, given the one-step earlier permutation π_i is known. Here, the information is measured in bits.

We first consider a simple case of m=2. There only exist 2 permutations: 01

and 10, where 01 indicates the increasing motif, and 10 indicates the decreasing motif. There can be 4 transition possibilities between these 2 permutations (Figure 1): (i) 01 → 01, (ii) 01 → 10, (iii)10 → 01, and (iv)10 → 10. Among these 4 types of transition, (i) and (iv) 10→10. Among these 4 types of transition, (i) and (iv) refer to no states change, while (ii) and (iii) refer to changing states. In Figure 1, we show the permutations transition in two-dimensional space. Below the diagonal line, x_t is always larger than x_{t+1}, and the permutation is 10. Whereas above the diagonal line, x_{t+1} is always larger than x_t, and the permutation is 01. For other m, the m-dimensional space will be divided into m! sub-spaces, and there can be $(m!)^2$ transition possibilities between each pair of permutations.

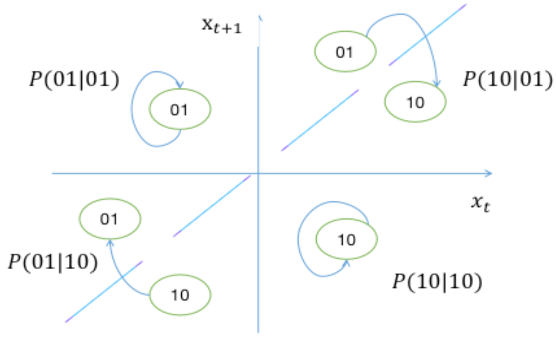


Figure 1: The 4 types of permutations transition between two permutations 01 and 10. Blow the dashed diagonal line, $\boldsymbol{x_t}$ is always larger than $\boldsymbol{x_{t+1}}$, and the permutation is 10. Above the dashed diagonal line,$\boldsymbol{x_{t+1}}$ is always larger than $\boldsymbol{x_t}$, and the permutation is 01.

2.3. Properties of the PTE method

we explore several properties of the PTE method.

1. The range of PTE.

By definition, T_m is always non-negative, i.e.$T_m > 0$. When the dynamical complexity is lowest, PTE is about 0. It means that there is very few permutations change. This is the case when the time series are persistently increasing or decreasing. On the other hand, PTE attains to its maximum value when all possible states changes are uniformly distributed. In such a case,$P(\pi_i) = 1/m!$, $P(\pi_i, \pi_j) = 1/(m!)^2$, $P(\pi_j|\pi_i) = 1/(m!)$, so the maximum value of the PTE is $2\log_2(m!)$ We can normalize the PTE by: $0 \le T_m/[2log_2(\mathrm{m}!)]$

2. Finite size effects

PE has a relatively low requirement on the data length, as $m! < T$ is appropriate. But it is not the same case for the PTE, where $m! \ll T$ is required. Therefore, it's suitable to set $m = 2,3,4,5$ for PTE. In Figure 2, we generate a series of Gaussian white noise with length $T = 10^6$ and set $m = 2, \ldots, 7$, respectively. The theoretical PTE results are obtained by $T_m = 2log_2\,(m!)$, compared with the empirical PTE values. When $m = 2,3,4,5$ the theoretical and empirical results coincide well. Deviations between the theoretical and empirical values first emerge at $m = 6$, and enlarge at $m = 7$, due to the finite size effects. Obviously, the deviations will postpone if the data length T enlarges.

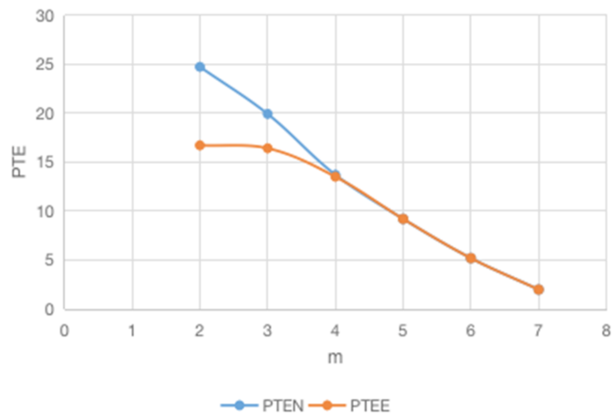


Figure 2: The theoretical PTE values $T_m = 2log_2(m!)$ with the blue color of the Gaussian white noise with data length $N = 10^6$, compared with the empirical PTE results with the orange color. The deviations between the theoretical and empirical results start at m=6 and enlarge at m=7 due to the finite size effects.

In many cases, $(5!)^2 = 14400$ is still too large for the length of time series. Here we present a strategy to tackle it.

We utilize the multiscale analysis of permutation transition entropy. A coarse-graining process is first applied before the estimation of PTE [9]. For the original time series $\{x_t\}, t = 1, \ldots, T,$ the non-overlapping coarse-grained time series at time scale s is given by:

$$y_j(s) = \frac{\sum_{i=(j-1)s+1}^{js} x_i}{s}, i \le j \le {}^{T}/_{s} \quad (7)$$

On scale one, the time series y is simply the original time series. The length of the coarse-grained time series is equal to the length of the original time series divided by the scale factor, s. Then, the PTE is estimated for the coarse-grained time series.

If the sampling frequency of the original time series is f, the sampling frequency of the sequence of permutations would be roughly f/m. Here the sequence of permutations refers to a series of discrete symbols of π projected from the state vectors. For a small data length T, we cannot set m large due to the finite size effects. So we can first down-sampling the original time series by the multiscale analysis to f/s. Then we set a smaller m^*, to make sure m $m \approx s \times m^*$ although the information reflected by PTE at the embedding. dimension m is not completely equivalent to the information reflected by multiscale PTE on scale s at the embedding dimension m^*, the multiscale analysis is still a promising approach to minimize the finite size effects [29].

3. False detection of permutations transition

If we cannot properly set the time delay τ in the phase space reconstruction, spurious permutations transition may arise. For example, for a series of the Gaussian white noise, the first state vector is $X_1 = [x_1, x_2, x_3]$ at $m = 3$ and $\tau = 1$. The second state vector is $X_2 = [x_2, x_3, x_4]$. If the permutation of X_1 is 012, which means that $x_1 < x_2 < x_3$, the permutation for X_2 cannot be 210, 102, or 201, since $x_2 < x_3$. Therefore, a false detection of the permutations transition may emerge due to artificially introducing connections between neighboring state vectors. [29] also mentions the similar case. There exist two approaches to solve this issue.

- The first approach is to set $\tau \geq 2$ for any m. Take $\tau \geq 2$ for example. The first state vector is $X_1 = [x_1, x_3, x_5]$, and the second state vector is $X_2 = [x_2, x_4, x_6]$. There is no artificial connection between X_1 and X_2. Although $X_3 = [x_3, x_5, x_7]$, and X_1 and X_3 both contain x_3, x_5, it is acceptable since we only consider the one-step permutations transition from $\pi(X_1)$ to $\pi(X_2)$.
- The second approach is to set $\tau = 1$, while use a time-delayed permutation transition [29,30]. For m=3, $X_1 = [x_1, x_2, x_3]$, $X_2 = [x_2, x_3, x_4]$, $X_3 = [x_3, x_4, x_5]$, and $X_4 = [x_4, x_5, x_6]$. Instead of incorporating the permutation transition from $\pi(X_1)$ to $\pi(X_2)$, we analyze the permutation transition from $\pi(X_1)$ to $\pi(X_4)$, then $\pi(X_2)$ to $\pi(X_5)$, and so on. That is to say, we introduce a transition delay, equal to the embedding dimension. The transition delay makes sure that the neighboring permutations are projected from state vectors that contain no overlapping data. For other m, we can derive the same conclusion.

3. Numerical Experiment and Analysis

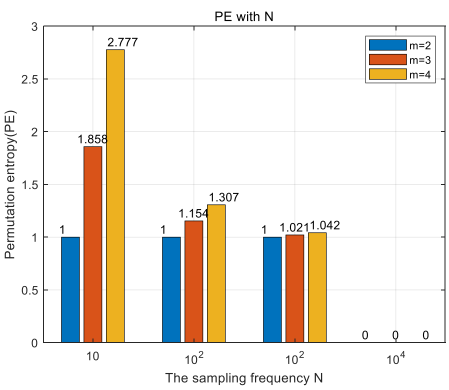
3.1. Artificial time series analysis

For a purely random time series, the theoretical PTE is $T_m = 2\log_2 m!$, which is the upper bound. For a persistently increasing (or decreasing) time series, there is only one type of permutation, and the PTE is $T_m = 0$, corresponding to the lower bound. For many other cases, $0 < H_m < 2log_2(m!)$.

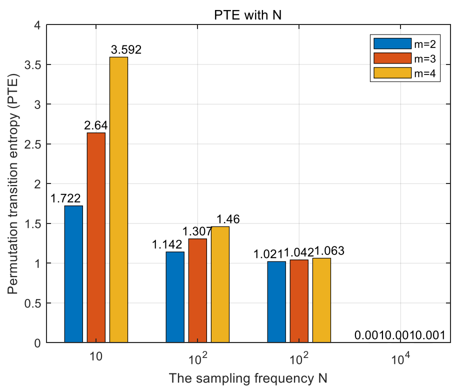
We first consider a simple periodic time series:

$$x_t = A_1 \sin\big(2\pi\phi_1(t)\big) \tag{8}$$

, where $t = 1, \ldots, T$, A_1 is the amplitude, $2\pi\phi_1(t)$ is the phase function, and $1/\phi_1'(t)$ is the sampling frequency. The dynamical complexity is dependent on the sampling frequency $1/\phi_1'(t)$, but is independent on the amplitude A_1. Generally, is we enlarge the sampling frequency, the dynamical complexity would decrease. An extreme case is to set the sampling frequency infinitely large, then at most time the state vectors turn out to have an increasing (or decreasing) trend. The absence of the permutations transition leads to a low degree of dynamical complexity. In Figure 3, we set $\phi_1(t) = t/N$, and N=10, 10^2, 10^3, 10^4, respectively. At a fixed m, the values of PTE, as expected, decrease with the increase of *N*. The values of PTE increase with the increase of m, since more states transition are likely to appear. In most cases, the PTE is much smaller than that of the random time series.



(**a**)



(**b**)

Figure 3: The values of PE and PTE with different sampling frequencies N and embedding dimensions m. **(a)**The values of PE for sinusoidal signal with the sampling frequency $N = 10,\ 10^2,\ 10^3,\ 10^4$, respectively. The PT decrease with the increase of N, accompanied by lower complexity. **(b)**The values of PTE for sinusoidal signal with the sampling frequency $N = 10, 10^2, 10^3, 10^4$,

respectively. The PTE decrease with the increase of N, accompanied by lower dynamical complexity. (Note: The values of PTE of m=2, 3, 4 are 5.77×10^{-4}, ,5.77×10^{-4}, 5.78×10^{-4} when $N = 10^4$.)

Next, we use the PTE method to analyze the logistic map. It is a polynomial mapping of degree 2:

$$x_{t+1} = \mu x_t(1 - x_t) \tag{9}$$

For $\forall t, x_t \in [0,1]$, which can be used to represent the ration of existing population to the maximum possible population in ecology. The values of interest for the parameter μ are those in the interval [0,4]. Complex, chaotic behavior can arise from this very simple non-linear dynamical equation. It was indicated that most values of μ beyond 3.56995 exhibit chaotic behavior. Here we set μ = 3.6,3.7,3.8,3.9 respectively, and set the data length T=10^5. The initial value of x_0 is set to 0.5. By changing the parameter μ, different types of behaviors can be observed.

In Table 1, we show the number of different permutations transition of the logistic map at $\mu = 3.6, 3.7, 3.8, 3.9$, respectively. We also give the number of permutations transition of the purely random time series with the same data length. The embedding dimension m is set to 2,3,4,5. The time delay is $\tau = 2$. The theoretical number of the purely random time series is $(m!)^2$. For the empirical data, there exist forbidden patterns [10]; it means that many permutations transition never arise, which is related to the temporal structure of the logistic map. The number of different permutations transition thus be smaller than $(m!)^2$. When $m = 2$, the number of different permutations transition is always 4 for each case. But when $m = 3$,4,5, the number for the random time series is much larger than that for the logistic map. Moreover, the number of different permutations transition increases with the increase of μ.

Table 1: The number of different permutations transition for μ = 3.6, 3.7, 3.8, 3.9 of the logistic map and for the random time series with different, respectively. The embedding dimension m is set 2, 3, 4, 5. When $m = 5$, the number for random time series is 14379, which is a bit smaller than the theoretical number of 14400 due to the finite size effects.

The parameter μ	$m = 2$	$m = 3$	$m = 4$	$m = 5$
$\mu = 3.6$	4	12	16	30
$\mu = 3.7$	4	15	38	99
$\mu = 3.8$	4	16	48	148
$\mu = 3.9$	4	18	62	319
Random	4	36	569	14379

In Table 2, we give the PTE values of the logistic map and the purely random time series, respectively. There is no significance difference among these series at $m = 2$. While for $m = 3$,4,5, the PTE values of the logistic map are significantly smaller that those of the random time series. Therefore, it indicates that the dynamical complexity of the logistic map is smaller than that of the purely random time series. Since there exists temporal structure of the logistic map to make the series a bit regular. It also manifests that the PTE can be used to effectively distinguish the chaotic behavior from the random behavior [31]. Moreover, with the increase of μ, the series of the logistic map behaves more random, as the PTE values become larger.

Table 2: The values of PTE for μ = 3.6,3.7,3.8,3.9 of the logistic map and for the random time series, respectively. The embedding dimension *m* is set to 2, 3, 4, 5. The theoretical values of PTE for random time series are 1.91, 4.09, 6.51, 9.16, respectively.

The parameter μ	$m = 2$	$m = 3$	$m = 4$	$m = 5$

$\mu = 3.6$	1.000	2.124	1.251	3.345
$\mu = 3.7$	1.238	2.544	3.004	3.925
$\mu = 3.8$	1.424	2.707	3.244	4.121
$\mu = 3.9$	1.501	2.793	4.056	4.977
Random	1.919	4.085	6.506	9.156

3.2. Stock index time series analysis

The study on the predictability of stock time series have attracted the close attention of researchers for many decades. If the efficient market hypothesis (EMH) is of some relevance to reality, then a market will be very unpredictable due to its capability to instantly digest any new information. However, many recent studies have shown that long-range persistent auto-correlations exist in the stock time series of emerging markets, e.g.by the Hurst exponent $H > 0.5$ [32-34]. It explains a fact that the memory of data may exist between neighboring values. Furthermore, the strength of memory is closely related to the predictability of the stock time series: the stronger memory of the series, the more likely to be predicted. For the trading of stocks, the investigators concern much on the stock price going up or down. The fluctuation of stock price can be reflected by the permutation of neighboring prices. For example, the permutation of 01 represents the stock price rising up, while 10 represents the stock price falling down. Moreover, 012 indicates that the stock price successively rising up at two sampling points e.g.two days, and so on.

Then, two classes of questions may arise: (1) Which permutation occurs the most frequently? Is there a dominant permutation? How much information is contained in all permutations? (2) Which permutation is likely to be followed by another permutation? How much information is contained in all the permutations transition? The first class of questions can be tackled by the PE, but not the second one. Here we try to use the PTE to solve the second class of questions. Here we try to use the PTE to solve the second class of questions.

In finance, there exist two popular investment strategies: one is related to the momentum effect; another one is the mean reversion effect. Momentum effect is the empirically observed tendency for rising asset prices to rise further, and falling prices to keep falling. For instance, it was shown that stocks with strong past performance continue to outperform stocks with poor past performance in the next period with an average excess return of about 1% per month [35-37] documented that strategies which buy stocks that have performed well in the past and sell stocks that have performed poorly in the past generate significant positive returns over 3- to 12- month holding periods. On the contrary, mean reversion is a theory used in finance that suggests that asset prices and historical returns eventually return back to the long-run mean or average level of the entire data set. The thought is that any price that strays far from the long-term norm will again return, reverting to its understood state. This theory has led to many investing strategies from stock trading to options pricing [38, 39]. Here, we would like to use the PTE method to judge the existence of the existence of the momentum effect or the mean reversion effect in the emerging markets.

We analyze the NASDAQ Composite Index in the U.S. The daily closing price data of this index is studied, which spans from January 1, 2000 to December 30, 2016 (T=4278). The logarithm change of volume is given by:

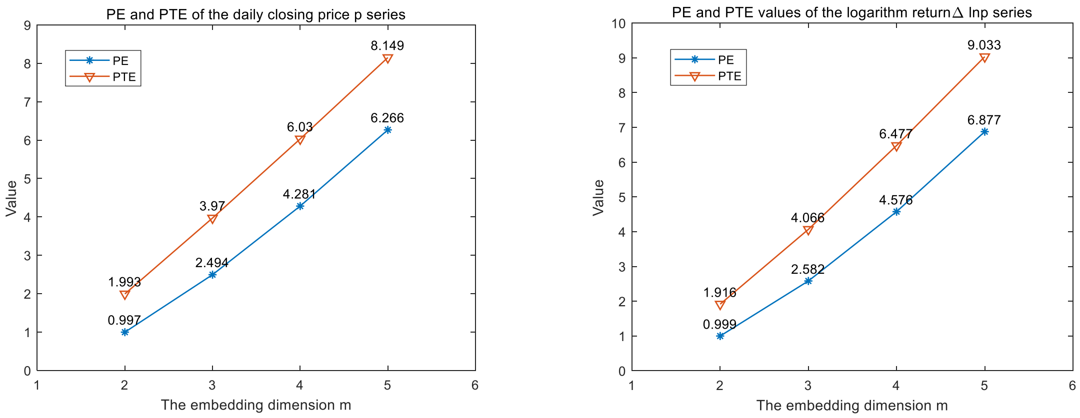
$$\Delta \ln(\mathrm{p}) = \ln(p_t) - \ln(p_{t-1}) \tag{10}$$

3.2.1. PE analysis of NASDAQ Composite Index

We first apply PE to estimate the complexity of p, $\Delta\, ln\,(p)$ in the time series of NASDAQ Composite Index. At the embedding dimension $m = 2$, the occurrence probability of 01 (describing the increase of stock price) and the probability of 10 (describing the decrease of stock price) are almost equal. Thus, the PE values at $m = 2$ are equal to that of the purely random time series. It indicates that to predict the increase or decrease of the closing price based on the closing price one day earlier is impossible. Therefore, as the answer of the first class of questions, there exists no dominant permutation which occurs much more frequent than other permutations. The information contained in all permutations of NASDAQ Composite Index series are very close to that of the purely random series. In other words, we cannot oppose the EMH with sufficient reasons by the PE method.

3.2.2. PTE analysis of NASDAQ Composite Index

We further apply the PTE method to estimate the dynamical complexity of p, $\Delta\ln(\mathrm{p})$ In NASDAQ Composite Index time series.



(a) (b)

Figure 4: **(a)** PE and PTE values of the daily closing price p series. For comparison, the theoretical PTE values of purely random time series are 2.00, 5.17, 9.17, 13.81, respectively. **(b)** PE and PTE values of the logarithm return $\Delta\ln(\mathrm{p})$ series. For comparison, the theoretical PTE values of purely random time series are 2.00, 5.17, 9.17, 13.81, respectively.

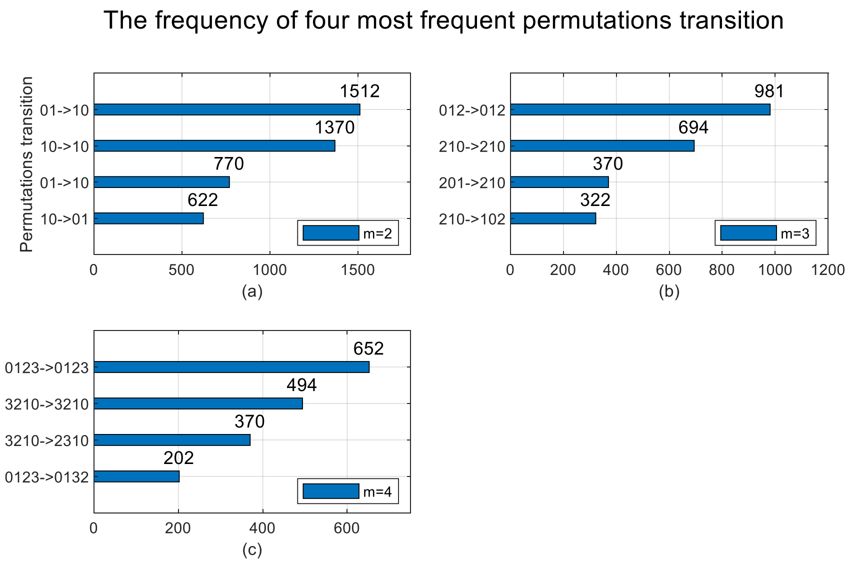


Figure 5: The four most frequent permutations transition for the daily closing price p in NASDAQ Composite Index, where the embedding dimension is set to m=2,3,4, respectively.

For the daily closing price series p at $m = 2$, the occurrence frequencies of $01 \rightarrow 01$(occurrence frequency: 1512) and $10 \rightarrow 10$(occurrence frequency:1370) are much larger than those of $01 \rightarrow 10$(occurrence frequency:770) and $10 \rightarrow 01$(occurrence frequency:622). Moreover, at m=3, the most frequent permutations transitions are $012 \rightarrow 012$(occurrence frequency:981). While the occurrence frequency of $012 \rightarrow 210$ is only 19, and that of $210 \rightarrow 012$ is only 23. For more details, please see Figure 5. It strongly indicates that the momentum effect is more prominent than the mean reversion effect for the daily closing price series in NASDAQ Composite Index.

The logarithm return $\Delta \ln(\mathrm{p})$ is much more random than the daily closing price series p, and hence with higher dynamical complexities. The PTE values for the logarithm return series is approximate to those of the purely random series.

From the analysis of PTE, we reject the EMH in NASDAQ Composite Index, since the momentum effect arises in the daily closing price series. It manifests that by PTE we can reveal new information which is not conveyed by PE.

4. Results

Firstly, we adopt a coarse-graining process to the original time series$\{x_t\}, t = 1, \ldots, T$.

$$y_j(s) = s^{-1} \sum_{i=(j-1)s+1}^{js} x_i \,, i \le j \le \frac{T}{s} \tag{11}$$

We analyze the cases when the time interval s=3,4, with different time delay m=2,3,4.

- When we are considering the transition of non-overlapping permutations, the value of PTE at s=3 are greater than the PTE at s=4. Moreover, the value of PE at s=3 are greater than the PE at s=4, which means that with lower time scale, the uncertainty of the daily closing price series is higher. Thus the information contained in the time series are fewer, and the predictability is lower.
- Under different time-scales, the momentum effect is more obvious in the daily closing price series with higher time-scale. For example, when the embedding dimension is 2 and the time-scale is 3, the occurrence frequency of 01→01 is 1853, the occurrence frequency of 10→10 is 1476. However, when the embedding dimension is 2 and the time-scale is 4, the occurrence frequency of 01→01 is 1976, and the occurrence frequency of 10→10 is 1534. So we can conclude that after the process of coarse-grinning, the certainty of the closing price time series increases. Furthermore, with higher time-scale, the certainty is stronger, and the predictability is higher.

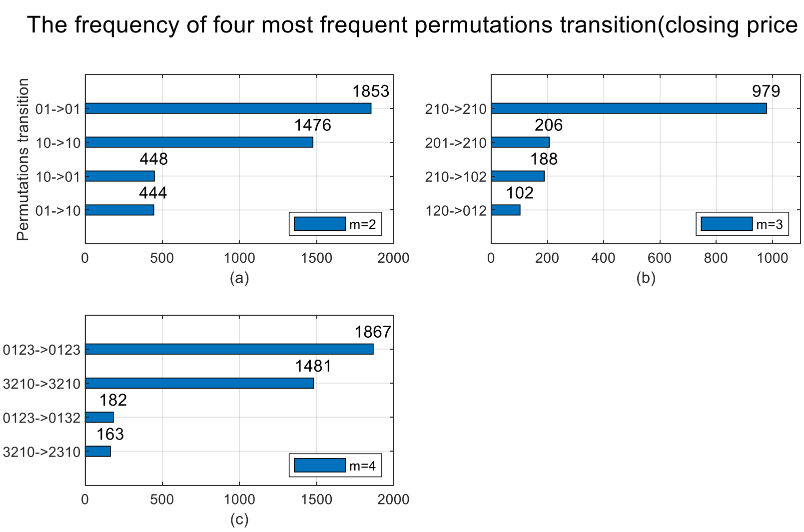


Figure 6: The four most frequent permutations transition for the daily closing price p in NASDAQ Composite Index under the time-scale 3, where the embedding dimension is set to m=2,3,4, respectively.

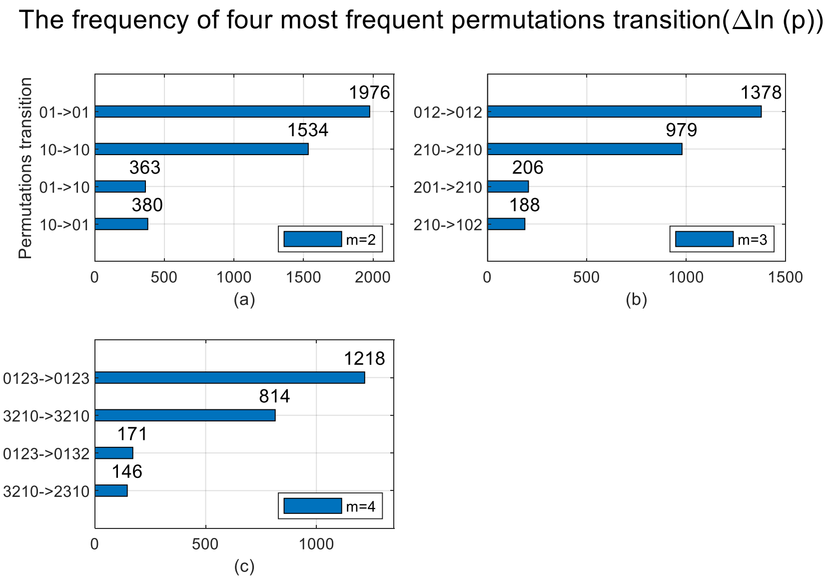
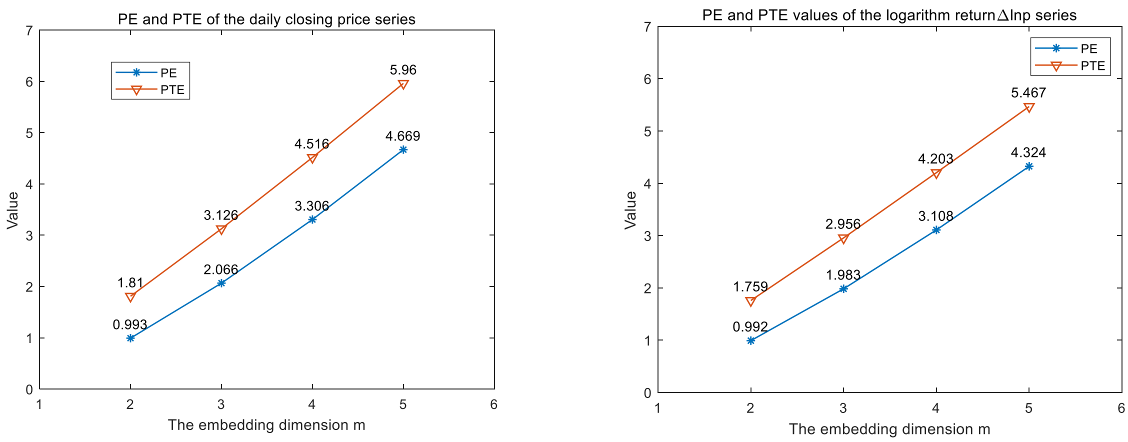


Figure 7: The four most frequent permutations transition for the logarithm return $\Delta \ln(p)$ in NASDAQ Composite Index under the time-scale 3, where the embedding dimension is set to m=2,3,4, respectively.



(**a**) (**b**)

Figure 8: **(a)** The values of PE and PTE of the daily closing price series under the time-scale of 3. **(b)**The values of PE and PTE of the the logarithm return $\Delta \ln(p)$ under the time-scale of 4.

5. Conclusions

In this paper, we propose the PTE method to characterize the dynamical complexity of a signal time series. The PTE incorporates the permutation analysis and the conditional entropy, which can describe the diversity of the forthcoming events in the presence of the current events. The upper and lower bounds of the PTE, the finite size effects, and any false detection of permutations transition are explored for the PTE method. The PTE is then used to analyze the periodic time series, and the logistic map. It turns out that the PTE method can be used to distinguish the chaotic behavior from the random behavior.

In the empirical analysis, the NASDAQ Composite Index is analyzed by the PTE method, by comparing with the PE method. The PTE method can solve a class of questions that are related to the dynamical complexity, while the PE method cannot. The momentum effect is found in the daily closing price of NASDAQ Composite Index. But the logarithm return behave very similar to the purely random time series. Furthermore, we utilized multiscale analysis towards the method of PTE, it turns out that under the same embedding dimension, the series with higher time-scale are more deterministic and represents more obvious momentum effect. The method of PTE requires no assumption on the time series; it can be applied for non-stationary and nonlinear time series analysis. Therefore, the PTE method can be further applied to more areas.

6. Patents

Author Contributions: Conceptualization, Xiaowei Mao. and Chaoquan Jiang.; methodology, Xiaowei Mao.; software, Zhigang Kou.; validation, Zhigang Kou, Xiaowei Mao; formal analysis, Xiaowei Mao and Chaoquan Jiang; investigation, Xiaowei Mao and Zhigang Kou; resources, Xiaowei Mao, Zhigang Kou; data curation, Zhigang Kou.; writing—original draft preparation, Xiaowei Mao.; writing—review and editing, Chaoquan Jiang and Lizhang Chen; visualization, Xiaowei Mao, chaoquan Jiang; supervision, Xiaowei Mao, Lizhang Chen; project administration, Xiaowei Mao; All authors have read and agreed to the published version of the manuscript.

Acknowledgments: This research supported by Beijing Jiaotong University Training Program of Innovation and Entrepreneurship for Undergraduates, Beijing Training Program of Innovation and Entrepreneurship for Undergraduates and National Training Program of Innovation and Entrepreneurship for Undergraduates.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. S. Lloyd, Measures of complexity: a non-exhaustive list, IEEE Control Systems Magazine 21 (4) (2001) 7–8.
2. W. Chen, J. Zhuang, W. Yu, Z. Wang, Measuring complexity using fuzzy, and sampen, Medical Engineering & Physics 31 (1) (2009) 61–68.
3. D.Ke, Unifying complexity and information, Scientific Reports 3(4)(2013) 1585.
4. J. S. Richman, J. R. Moorman, Physiological time-series analysis using approximate entropy and sample entropy, American Journal of Physiology Heart & Circulatory Physiology 278 (6) (2000) H2039.
5. X. Zhao, P. Shang, J. Huang, Permutation complexity and dependence measures of time series, EPL 102 (4) (2013) 40005.
6. T.Schreiber, Measuring information transfer, Physical Review Letters85(2) (2000) 461–464.
7. M. Lungarella, K. Ishiguro, Y. Kuniyoshi, N. Otsu, Methods for quantifying the causal structure of bivariate time series, International journal of bifurcation and chaos 17 (03) (2007) 903–921.
8. M.Jiang, X.Gao, H.An, H.Li, B.Sun, Reconstructing complex network for characterizing the time-varying causality evolution behavior of multivariate time series, Scientific Reports 7 (1) (2017) 10486.
9. M. D. Costa, A. L. Goldberger, C. K. Peng, Multiscale entropy analysis of complex physiologic time series, Physical Review Letters 89 (6) (2002) 068102.
10. L.Zunino, M. Zanin, Complexity entropy causality plane: A useful approach to quantify the stock market inefficiency, Physica A 389 (9) (2010) 1891–1901.
11. C.K.Peng, S.V.Buldyrev, A.L.Goldberger, S.Havlin, F.Sciortino, M.Simons, H. E. Stanley, Long-range correlations in nucleotide sequences, Nature 356 (6365) (1992) 168.
12. A.Wolf, J.B.Swift, H.L.Swinney, J.A.Vastano, Determining lyapunov exponents from a time series, Physica D: Nonlinear Phenomena 16 (3) (1985) 285–317.
13. B. Podobnik, D. Horvatic, A. M. Petersen, H. E. Stanley, Cross-correlations between volume change and price change, Proceedings of the National Academy of Sciences 106 (52) (2009) 22079–22084.
14. A. Forsythe, M. Nadal, N. Sheehy, C. J. Cela-Conde, M. Sawey, Predicting beauty: fractal dimension and visual complexity in art, British journal of psychology 102 (1) (2011) 49–70.
15. S.M.Pincus, Approximate entropy as a measure of system complexity, Proceedings of the National Academy of Sciences 88 (6) (1991) 2297–2301.

16. R. J. Balvers, Y. Wu, Momentum and mean reversion across national equity markets, Journal of Empirical Finance 13 (1) (2006) 24–48.
17. Xia,P.Shang,J.Wang,W.Shi, Permutation and weighted-permutation entropy analysis for the complexity of nonlinear time series, Communications in Nonlinear Science and Numerical Simulation 31 (1-3) (2016) 60–68.
18. X. Zhao, P. Shang, Measuring the uncertainty of coupling, EPL 110 (6) (2015) 60007.
19. X. Zhao, P. Shang, J. Wang, Measuring the asymmetric contributions of individual subsystems, Nonlinear Dynamics 78 (2) (2014) 1149–1158.
20. C.Bandt, B.Pompe, Permutation entropy: a natural complexity measure for time series, Physical Review Letters 88 (17) (2002) 174102.
21. C. Y., T. Wen-Wen, J. B. Gao, V. A. Protopopescu, L. M. Hively, Detecting dynamical changes in time series using the permutation entropy, Physical Review E 70 (4 Pt 2) (2004) 046217.
22. R. K. Y. Low, E. Tan, The role of analyst forecasts in the momentum effect, International Review of Financial Analysis 48 (2016) 67–84.
23. Y. Yin, P. Shang, Modified DFA and DCCA approach for quantifying the multiscale correlation structure of financial markets, Physica A 392 (24) (2013) 6442–6457.
24. R. J. Balvers, Y. Wu, Momentum and mean reversion across national equity markets, Journal of Empirical Finance 13 (1) (2006) 24–48.
25. F. Traversaro, N. Ciarrocchi, F. P. Cattaneo, F. O. Redelico, Comparing dif- ferent approaches to compute permutation entropy with coarse time series, Physica A 513 (2019) 635–643.
26. J. Gao, F. Liu, J. Zhang, J. Hu, Y. Cao, Information entropy as a basic build- ing block of complexity theory, Entropy 15 (9) (2013) 3396–3418.
27. F. Takens, Dynamical Systems and Turbulence, Springer, Berlin,, 1981.
28. L. Zunino, D. G. Pe´rez, M. T. Mart´ın, M. Garavaglia, A. Plastino, O. A. Rosso, Permutation entropy of fractional brownian motion and fractional gaussian noise, Physics Letters A 372 (27-28) (2008) 4768–4774.
29. X. Zhao, Y. Sun, X. Li, P. Shang, Multiscale transfer entropy: measuring information transfer on multiple time scales, Communications in Nonlinear Science and Numerical Simulation 62 (2018) 202–212.
30. S. D. Wu, C. W. Wu, S. G. Lin, K. Y. Lee, C. K. Peng, Analysis of complex time series using refined composite multiscale entropy, Physics Letters A 378 (20) (2014) 1369–1374.
31. O. A. Rosso, H. A. Larrondo, M. T. Martin, A. Plastino, M. A. Fuentes, Distinguishing noise from chaos, Physical Review Letters 99 (15) (2007) 154102.
32. C. K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, A. L. Goldberger, Mosaic organization of DNA nucleotides, Physical Review E 49 (2) (1994) 1685.
33. Y. Yin, P. Shang, Modified DFA and DCCA approach for quantifying the multiscale correlation structure of financial markets, Physica A 392 (24) (2013) 6442–6457.
34. H. Xiong, P. Shang, Detrended fluctuation analysis of multivariate time se- ries, Communications in Nonlinear Science and Numerical Simulation 42 (2017) 12–21.
35. R. K. Y. Low, E. Tan, The role of analyst forecasts in the momentum effect, International Review of Financial Analysis 48 (2016) 67–84.
36. N. Jegadeesh, S. Titman, Returns to buying winners and selling losers: Im- plications for stock market efficiency, The Journal of Finance 48 (1) (1993) 65–91.
37. N. Jegadeesh, S. Titman, Profitability of momentum strategies: An evalua- tion of alternative explanations, Journal of Finance 56 (2) (2001) 699–720.
38. M. H. Miller, J. Muthuswamy, R. E. Whaley, Mean reversion of standard & poor's 500 index basis changes: Arbitrage-induced or statistical illusion?, The Journal of Finance 49 (2) (1994) 479–513.
39. R. J. Balvers, Y. Wu, Momentum and mean reversion across national equity markets, Journal of Empirical Finance 13 (1) (2006) 24–48.
40. B. Podobnik, D. Horvatic, A. M. Petersen, H. E. Stanley, Cross-correlations between volume change and price change, Proceedings of the National A- cademy of Sciences 106 (52) (2009) 22079–22084.