Physica A 633 (2024) 129398

Contents lists available at ScienceDirect

Physica A

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

Identify causality by multi-scale structural complexity

Ping Wang [a], Changgui Gu [b,*], Huijie Yang [b], Haiying Wang [b]

[a] School of Science, Jiangsu University of Science and Technology, Zhenjiang, Jiangsu, 212114, China
[b] Department of Systems Science, Business School, University of Shanghai for Science and Technology, Shanghai, 200093, China

ARTICLE INFO

Keywords:
Causality
Multi-scale structural complexity
Synthetic data
Empirical data

ABSTRACT

Identifying causality between variables is essential to reveal their interactions. In previous studies, causality between variables has been captured at specific scales. However, it was not taken into account the difference in causality between different scales. In other words, it remains an open question about which scale of causality contributes most to correct causality. In this paper, a multi-scale structural complexity method of bivariate time series was developed from the recent advances in measuring the complexity of visual images. This method can identify the direction and strength of causality between variables by total complexity. In addition, depending on the contribution from the local complexity of each scale to the total complexity, it determines which scaling scale(i.e., levels of resolution) is responsible for causality. Our results indicated that local complexity at the medium scale contributes the most to total complexity. More interestingly, this conclusion holds not just for synthetic data but also for empirical data. Consequently, our finding provides not only a novel perspective to reveal the causality between variables from the perspective of complexity, but also a reference to choose an appropriate scale to study the relationship between variables.

1. Introduction

Causal interactions are fundamental underpinnings in various fields, such as ecological systems [1], financial markets [2–4], and physiological systems [5,6]. Identifying correct causality between the dynamical variables described by the form of the time series provides the possibility to study the internal dynamics of the system [7]. Several classical methods have been developed to infer causality between variables, such as the Granger causality method which can demonstrate linear causality between separable variables [8–10], the entropy method that can detect causality via the flow of information between variables [11–13], and the convergent cross mapping (CCM) method, as a supplement to the Granger causality method, which can distinguish causality from the correlation in inseparable systems [1]. By all above methods, causality between variables has been captured at specific scale-size. However, it was not taken into account the difference in causality at different scales.

Several previous studies have shown that the temporal scale-size affects the inference of causality [14,15]. If the scale is too large, the rules of causality between variables is obscured; while if the scale is too small, it is more vulnerable to noise, and both the complexity and the research cost are higher. Although research has indicated that sampling frequency can adjust the scale-size of the study, a study with an inappropriate sampling frequency may have spurious effects on testing causality [16,17]. Consequently, the difference in causality at different scales should be considered (i.e., the causality at multiple scales). In other words, it remains an open question about which scale of causality contributes most to correct causality.

Recently, a multi-scale structural complexity method (namely, MSC) was proposed by Bagrov et al. to characterize complexity at multiple spatial scales for image data [18]. The MSC method can accurately identify different phase transitions and provide

* Corresponding author.
E-mail address: gu_changgui@163.com (C. Gu).

https://doi.org/10.1016/j.physa.2023.129398
Received 19 September 2023; Received in revised form 4 November 2023
Available online 24 November 2023

information about the dynamics of non-equilibrium systems [18,19]. The authors in Ref. [20] exploited the MSC method to develop a multi-scale structural complexity method based solely on univariate time series (namely, S-MSC). In addition to accurately reflecting the dynamic characteristics of a system, the S-MSC method can also identify the local complexity of a system at multiple scales. Moreover, it can also indicate how the complexity of the system changes with scale-size. Accordingly, the S-MSC method resolves the multi-scale problem ignored by traditional causal analysis ignores. It is applicable to univariate time series. In this article, we developed a novel method based on bivariate time series for measuring multi-scale causality between variables (namely, BS-MSC).

In this article, the validation of the BS-MSC method was performed on the synthetic time series (including the Logistic model and the Rössler–Lorenz model) as well as the empirical time series (i.e., the gait data). Based on this method, we can obtain the total complexity for characterizing causality between variables as well as the local complexity at each scaling scale. More interestingly, the total complexity varies with the delay time between bivariate time series, which determines the causality direction (because the reason always precedes the result). The following sections are organized as follows. In Section 2, we present (i) all of the models to generate the synthetic data of time series, as well as the empirical data; (ii) the description of BS-MSC method. Section 3 shows (i) the variation of total complexity for bivariate with the delay time; (ii) the variation of local complexity for bivariate with the delay time; Finally, Section 4 summarizes and discusses the results.

2. Data and methods

In this section, we introduced two models, including the Logistic model of bidirectionally coupled oscillators and the Rössler–Lorenz model of unidirectionally coupled oscillators. By the models, the synthetic time-series data for current study was generated. The length of the synthetic time series for each model is 50000 data points. In this section, the experimental data used in this study was presented, namely the gait data.

2.1. Description of data

2.1.1. The Logistic model of bidirectional coupling

The Logistic model is an early example of chaos created by the period-doubling process [21]. The dynamic behavior of this system strongly depends on the value of γ which determines whether the system is chaotic or periodic. The Logistic model of bidirectionally coupling oscillators, as in Refs. [1,22], was built by adding coupled terms to the Logistic system of each variate. The model is defined as,

$$\begin{cases} X_{t+1} = X_t(\gamma_x - \gamma_x X_t - aY_t), \\ Y_{t+1} = Y_t(\gamma_y - \gamma_y Y_t - bX_t), \end{cases} \tag{1}$$

where the parameters were $\gamma_x = 3.8$ and $\gamma_y = 3.9$, and initial conditions were $X(1) = 0.4$, $Y(1) = 0.2$. Herein, the coupled coefficient "a" represents the influence strength of variable Y on variable X, and the coupled coefficient "b" represents the influence strength of variable X on variable Y.

2.1.2. The Rössler–Lorenz model of unidirectional coupling

The Lorenz weather model [23], composed of three equations, is the first model by which both chaos and strange attractors were recognized. Its dynamic behavior is related to parameters d, e and f. The model is described as

$$\begin{cases} \dot{X} = d(-X + Y), \\ \dot{y} = eX - Y - XZ, \\ \dot{z} = XY - fZ. \end{cases} \tag{2}$$

The Rössler model was proposed by German scientist Rössler O. E. in 1976, which is composed of three nonlinear ordinary-differential-equations with parameters g, h and j [24–26]. Specifically, the equations are

$$\begin{cases} \dot{U} = -g(V + W), \\ \dot{V} = g(U + hV), \\ \dot{W} = g(h + W(U - j)). \end{cases} \tag{3}$$

Throughout this article, we set $d = 6$, $e = 28$, $f = \frac{8}{3}$, $g = 6$, $h = 0.2$ and $j = 5.7$ to ensure that the time series generated for the variables X, Y, Z, U, V and W are chaotic. To demonstrate the coupled relationship between Eqs. (2) and (3), we consider the coupled Rössler–Lorenz system discussed by Le Van Quyen et al. [1,27]. The Rössler–Lorenz model is defined by the following differential equations,

$$\begin{cases} \dot{U} = -6(V + W), \\ \dot{V} = 6(U + 0.2V), \\ \dot{W} = 6(0.2 + W(U - 5.7)), \\ \dot{X} = 10(-X + Y), \\ \dot{Y} = 28X - Y - XZ + GV^2, \\ \dot{Z} = XY - \frac{8}{3}Z, \end{cases} \tag{4}$$

Table 1
ID numbers for the subjects.

Parkinson group						Healthy group			
Original ID	ID	Original ID	ID	Original ID	ID	Original ID	ID	Original ID	ID
GaPt03	Pt-01	GaPt15	Pt-11	GaPt25	Pt-21	GaCo01	Co-01	GaCo11	Co-11
GaPt04	Pt-02	GaPt16	Pt-12	GaPt26	Pt-22	GaCo02	Co-02	GaCo12	Co-12
GaPt05	Pt-03	GaPt17	Pt-13	GaPt27	Pt-23	GaCo03	Co-03	GaCo13	Co-13
GaPt06	Pt-04	GaPt18	Pt-14	GaPt28	Pt-24	GaCo04	Co-04	GaCo14	Co-14
GaPt07	Pt-05	GaPt19	Pt-15	GaPt29	Pt-25	GaCo05	Co-05	GaCo15	Co-15
GaPt08	Pt-06	GaPt20	Pt-16	GaPt30	Pt-26	GaCo06	Co-06	GaCo16	Co-16
GaPt09	Pt-07	GaPt21	Pt-17	GaPt31	Pt-27	GaCo07	Co-07	GaCo17	Co-17
GaPt12	Pt-08	GaPt22	Pt-18	GaPt32	Pt-28	GaCo08	Co-08	GaCo22	Co-18
GaPt13	Pt-09	GaPt23	Pt-19	GaPt33	Pt-29	GaCo09	Co-09		
GaPt14	Pt-10	GaPt24	Pt-20			GaCo10	Co-10		

where G is the strength of the unidirectional coupling between the two attractors. In other words, it represents the influence strength of the variable V on the variable Y. Note that the sampling interval in this paper was $\Delta t = 0.01$.

2.1.3. Gait data

The gait data can be downloaded free of charge from the famous website for physiology: https://physionet.org/content/gaitpdb. The data were obtained from each subject walking on a horizontal plane for approximately two minutes. Each foot was equipped with 8 sensors whose output was digitized and recorded at 100 samples per second, i.e., the length of each record is 12119 data points. In addition, the records also include two signals that reflect the sum of each foot's eight sensors. Consequently, each subject contains 18 records. In this article, our analysis is based on the sum of the eight sensor outputs for each foot. For more details about the data descriptions and brief analysis, please see Refs. [28–30].

In Table 1 we present a summary of all the samples considered in this paper. The subjects were divided into two groups, i.e., the experimental group (Parkinson's group, abbreviated as Pt) and the control group (Healthy group, abbreviated as Co). The number of subjects in the two subgroups is 29 and 18, respectively. Here, "Original ID" represents the original ID of the subjects in the database, and "ID" represents the ID in this article of eliminating the missing subject ID.

2.2. Method

2.2.1. The BS-MSC method

Following recent advances in measuring the complexity of visual images [15,18], we developed a multi-scale structural complexity method ($BS-MSC$) that examines causality for bivariate time series. Naturally, the first challenge we faced is to convert bivariate time series into a matrix. In this article, it was solved by an approach based on the scatter plots and the correlation of bivariate, as in Ref. [20](note that the details are slightly different). Specifically, the steps are as follows.

Step 1. Transformation from bivariate time series (X and Y) to a matrix T_0.

Firstly, given the bivariate time series $X = (X(1), X(2), X(3), \dots, X(N))$, $Y = (Y(1), Y(2), Y(3), \dots, Y(N))$ and the delay time τ. We can obtained the truncated time series $X_1 = (X(1), X(2), X(3), \dots, X(N-\tau))$ and the truncated time series $Y_1 = (Y(1+\tau), Y(2+\tau), Y(3+\tau), \dots, Y(N))$, where N is the length of the time series.

Secondly, a coordinate system of two dimension was formed by mapping X and Y to a horizontal axis and a vertical axis, respectively. The investigated region is from X_{min} to X_{max} and from Y_{min} to Y_{max}, where "min" and "max" represents the minimum and maximum values of the time series X or Y. The range from X_{min} to X_{max} (or Y_{min} to Y_{max}) was divided equally into L parts, so that a total of L^2 grids can be generated.

Finally, we defined a matrix T_0 with L rows and L columns. The value of element $T_0(i,j)$ is equal to the mean value of all points in the $(i,j)_{th}$ grid. Please see Fig. S1 in Appendix A for an illustration.

Then, the matrix T_0 was analyzed at each scale in Step 2.

Step 2. Calculation of MSC for matrix T_0.

Firstly, the matrix for each scale was calculated. The matrix T_0 was covered by a module with a size of $\xi \times \xi$ (Note that in this article, we set ξ to its minimum value, i.e., $\xi = 2$, as increasing the value of ξ not only prolongs the computational time but also does not yield more useful information about the system. In order to demonstrate whether the value of ξ affects the main findings, we compared the results obtained by $\xi = 2$ (see Fig. 1(a)) and $\xi = 3$ (see Fig. S3 in Appendix C), respectively.). The matrix T_1 can be obtained by replacing the elements in each module with the average value of elements within the module. Next, the matrix T_1 was covered by a module with a size of $\xi^2 \times \xi^2$. The matrix T_2 can be obtained by replacing the elements in each module with the average value of elements within the module. By analogy, we can obtain the T_3, T_4, …, T_n, where $\xi^n \times \xi^n = L \times L$. The specific calculation formula of matrix T_k is as follows:

$$T_k(i,j) = \frac{1}{\xi^k \times \xi^k} \sum_{m=1}^{\xi^k} \sum_{l=1}^{\xi^k} T_{k-1}\left(\left\lfloor \frac{i-1}{\xi^k} \right\rfloor \cdot \xi^k + m, \left\lfloor \frac{j-1}{\xi^k} \right\rfloor \cdot \xi^k + l \right), \tag{5}$$

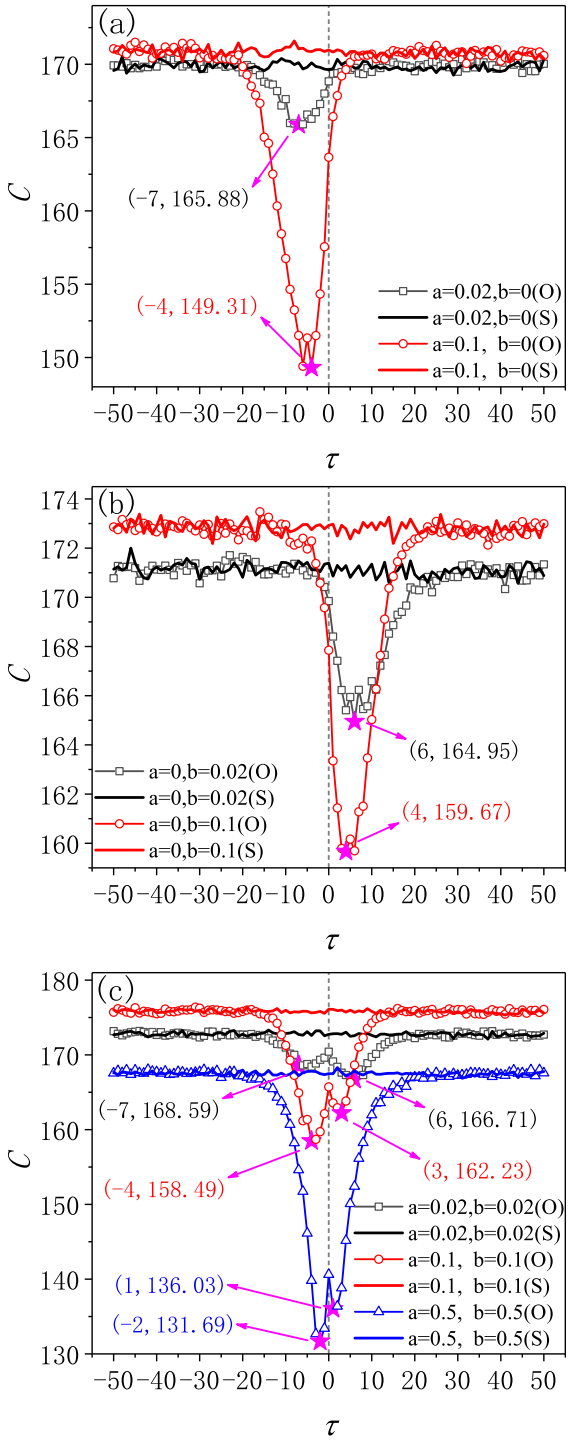

Fig. 1. The total complexity of the Logistic model (Eq. 2.1.1). (a) The total complexity for the parameters $a = 0.02, b = 0$ (black curves) and $a = 0.1, b = 0$ (red curves). (b) The total complexity for the parameters $a = 0, b = 0.02$ (black curves) and $a = 0, b = 0.1$ (red curves). (c) The total complexity for the parameters $a = b = 0.02$ (black curves), $a = b = 0.1$ (red curves) and $a = b = 0.5$ (blue curve). Here, hollow shapes and solid lines represent original sequences (namely, "O") and shuffled sequences (namely, "S"), respectively.

where the operator "$\lfloor\cdot\rfloor$" denotes rounding down and the integer k represents the scale number ($1 \le k \le \log_{\xi} L$). Figure S2 in Appendix B illustrates this step more intuitively with $n = 3$.

Secondly, the local complexity for each scale was calculated. The difference C_1 between T_0 and T_1 was calculated as the local complexity (i.e., the hierarchical complexity) for the first scale (i.e., $k = 1$). Next, the difference C_2 between T_1 and T_2 was calculated as the local complexity of the second scale (i.e., $k = 2$). By analogy, we can obtain the difference C_3, C_4, ..., C_n (i.e., $k = 3$, $k = 4$, ..., $k = n$). Specifically, the local complexity C_k of the matrix T_0 in k_{th} scale is defined as:

$$C_k = \left| \sum_{i=1}^{L} \sum_{j=1}^{L} O_k(i,j) \right|, \tag{6}$$

where

$$\begin{aligned} O_k &= -\frac{1}{2}\left(T_k - T_{k-1}\right) \cdot \left(T_k - T_{k-1}\right) \\ &= T_k \cdot T_{k-1} - \frac{1}{2}\left(T_k \cdot T_k - T_{k-1} \cdot T_{k-1}\right), \end{aligned}$$

where the operator "$\cdot$" represents dot multiplication between matrices.

Finally, the total complexity was calculated. According to Eqs. (5) and (6), the total complexity of the matrix T_0 is the sum of the local complexities, namely,

$$C = \sum_{k=1}^{\log_{\xi} L} C_k. \tag{7}$$

2.2.2. The relationship between causality and BS-MSC

According to the BS-MSC method, C represents the total complexity between the truncated time series $X_1 = (X(1), X(2), X(3), \ldots, X(N-\tau))$ and the truncated time series $Y_1 = (Y(1+\tau), Y(2+\tau), Y(3+\tau), \ldots, Y_N)$. Research has shown that the response variable is better at predicting the past values of the driving variable rather than future values [31], i.e., a reason always precedes a result. Hence, if $\tau > 0$, indicating that X precedes Y, it signifies the direction of causality as X driving Y; if $\tau < 0$, signifying that Y precedes X, it indicates the direction of causality as Y driving X. Hence, if $\tau > 0$, the value of C represents the degree of the variable X causing the variable Y; if $\tau < 0$, C represents the degree of the variable Y causing the variable X. In addition, the shuffled sequence is obtained by randomly disrupting the positions of the data points in the original sequence. To determine whether variable X and variable Y are causally related, C of the original sequence and C of the shuffled sequence are compared.

Assuming that X and Y are causally related, two aspects are required to be taken into account. First, if the value of the coupling coefficient (i.e., a or b for Eq. 2.1.1 and G for Eq. (4)) is different, the value of the complexity C is different with the focused value of τ. Specifically, the larger the value of C is, the weaker the causality between X and Y is, and vice versa. Therefore, the causality strength between X and Y can be determined by the value of C. Second, when the value of the coupling coefficient is the same, we are concerned about the value of τ with which the minimum complexity (i.e., C_{min}) is obtained. It makes sense to focus on the value of τ corresponding to C_{min} since C_{min} represents the strongest causality (i.e., C_{min} provides a better indication of causality between X and Y). If the value of τ corresponding to C_{min} is positive, it means that X causes Y; if the value of τ corresponding to C_{min} is negative, it means that Y causes X; if the value of τ corresponding to C_{min} is equal to 0, it indicates that X and Y occur simultaneously. There are enough reasons for us to believe that the direction of causality is mutual driving between X and Y, i.e., it means that X causes Y as well as Y causes X. Hence, the causality direction between X and Y can be identified by the value of τ corresponding to C_{min}.

3. Results

3.1. The Logistic model of bidirectional coupling

3.1.1. The influence of delay time on the total complexity

In this section, we presented the influence of delay time τ on the total complexity C for Eq. 2.1.1 with selected parameters (i.e., a and b). Three cases of values of a and b are examined, including $a \neq 0, b = 0$, $a = 0, b \neq 0$ and $a \neq 0, b \neq 0$ in Fig. 1.

The first case of $a \neq 0, b = 0$, including $a = 0.02, b = 0$ and $a = 0.1, b = 0$, is examined in Fig. 1(a). It is visible that the relationship of C to τ is parabolic-like with upward opening for the original sequence, but it is horizontally linear for the shuffled sequence, i.e., the influence of τ on C is negligible. In other words, causality exists between X and Y in both values of a. A trough exists for the original sequence, which value is $C_{min} = 165.88$ and $C_{min} = 149.31$ when $\tau = -7$ and $\tau = -4$ for $a = 0.02, b = 0$ and $a = 0.1, b = 0$, respectively. The trough values are much smaller than the complexity C for the shuffled sequence. Therefore, the variable Y causes the variable X because of $\tau < 0$. Additionally, C_{min} for $a = 0.1$ is smaller than C_{min} for $a = 0.02$, i.e., the causality strength for $a = 0.1$ is larger than the causality strength for $a = 0.02$.

The second case of $a = 0, b \neq 0$, including $a = 0, b = 0.02$ and $a = 0, b = 0.1$, is examined in Fig. 1(b). For the original sequence, the relationship between C and τ is parabolic-like with upward opening, but for the shuffled sequence, the relationship is horizontally linear, i.e., the influence of τ on C is negligible. Thus, it is true that causality exists between X and Y in both values of b. For the original sequence, a trough is also found at $C_{min} = 164.95$ and $C_{min} = 159.67$ when $\tau = 6$ and $\tau = 4$ for $a = 0, b = 0.02$ and

Table 2
The correlation coefficient r between total complexity and local complexity of the Logistic model (Eq. 2.1.1) for $a \neq 0, b = 0$.

The scale number k	$a = 0.02,\ b = 0$	$a = 0.1,\ b = 0$
	The correlation coefficient	The correlation coefficient
1	−0.85	−0.96
2	−0.87	−0.98
3	−0.90	−0.98
4	−0.75	−0.95
5	0.73	0.93
6	0.97	0.99
7	0.95	0.99
8	0.93	0.95
9	0.81	0.90
10	0.72	0.92
11	0.52	0.86
12	0.44	0.77

$a = 0, b = 0.1$, respectively. Consequently, X causes Y because of $\tau > 0$. Moreover, C_{min} for $b = 0.1$ is smaller than C_{min} for $b = 0.02$, i.e., the causality strength for $b = 0.1$ is larger than that for $b = 0.02$.

Finally, we study the last case of $a \neq 0, b \neq 0$ (including $a = b = 0.02$, $a = b = 0.1$ and $a = b = 0.5$) in Fig. 1(c). Similarly, it is evident that the relationship of C to τ is parabolic-like with upward opening for the original sequence, but it is horizontally linear for the shuffled sequence, i.e., the influence of τ on C is negligible. Hence, causality exists between X and Y in three values of a or b. Unlike the previous two cases, two troughs exists for the original sequence, which value is $C_{min_1} = 168.59, C_{min_2} = 166.71$, $C_{min_1} = 158.49, C_{min_2} = 162.23$ and $C_{min_1} = 131.69, C_{min_2} = 136.03$ when $\tau_1 = -7, \tau_2 = 6$, $\tau_1 = -4, \tau_2 = 3$ and $\tau_1 = -2, \tau_2 = 1$ for $a = b = 0.02$, $a = b = 0.1$ and $a = b = 0.5$, respectively. That is, for each original sequence, one trough corresponds to a negative τ(i.e., τ_1), while the other corresponds to a positive τ (i.e., τ_2). Thus, Y causes X because of $\tau_1 < 0$ as well as X causes Y because of $\tau_2 > 0$. Moreover, C_{min_1} (or C_{min_2}) follows a hierarchy, with C_{min_1} (or C_{min_2}) for $a = b = 0.5$ consistently lower than C_{min_1} (or C_{min_2}) for $a = b = 0.1$, which in turn is less than C_{min_1} (or C_{min_2}) for $a = b = 0.02$. In other words, the causality strength for $a = b = 0.5$ is greater than the causality strength for $a = b = 0.1$, which in turn is less than the causality strength for $a = b = 0.02$.

To sum up, for all three cases (i.e., $a \neq 0, b = 0$, $a = 0, b \neq 0$ and $a \neq 0, b \neq 0$), the simulation results are consistent with the values of coupling coefficient a and b. This consistency includes both consistency in causality direction and consistency in causality intensity. Consequently, the total complexity C can be utilized as an index to reflect the direction and strength of causality between variables for Eq. 2.1.1.

3.1.2. The influence of delay time on the local complexity

As we discussed in the previous part, the total complexity between variables X and Y reveals causality between them. So does the local complexity C_k, which is also based on the BS-MSC method, also capture causality between them? If C_k can, is the function of C_k for each scale number k the same or different? Furthermore, if different, which C_k has the most significant impact? In this part, we will study how C_k for each scale number k is varied with delay time τ. Likewise, three cases of values of a and b are examined, including $a \neq 0, b = 0$, $a = 0, b \neq 0$ and $a \neq 0, b \neq 0$ in Figs. 2–4.

Fig. 2 depicts the influence of delay time τ on local complexity C_k for the first case of $a \neq 0, b = 0$ with the scale number k from 1 to 12. The result illustrates that the relationship of C_k to τ is parabolic-like for the original sequence, but it is horizontally linear for the shuffled sequence, i.e., the influence of τ on C_k is negligible. Specifically, when $1 \leq k \leq 4$ (i.e., a relatively small scale), the parabolic-like for C_k is open downward, as opposed to a parabolic-like for C (see panels (a-d)). It can also be verified by the negative correlation coefficient between C_k and C in Table 2. The reason for opposite trends may be that the system is susceptible to noise interference when k is relatively small. Beyond that, we are unable to provide a more plausible explanation at present; when $k > 4$, the parabolic-like of C_k is open upward, i.e., the same as the parabolic-like of C (see panels (e–l)). In particular, when $k = 6$ or $k = 7$ (i.e., the medium scale), the trends of C_k and C are highly consistent. It can be verified by the maximum correlation coefficients obtained at $k = 6$ for $a = 0.02, b = 0$ and $k = 6$ or $k = 7$ for $a = 0.1, b = 0$, respectively, as shown by the underlined numbers in Table 2; when $k > 7$ (i.e., a relatively large scale), the correlation coefficient in Table 2 decreases with increasing k, indicating that C_k and C are less similar with increasing k.

Fig. 3 presents the influence of τ on C_k for the second case of $a = 0, b \neq 0$ with k from 1 to 12. Interestingly, the main results in Fig. 3 are essentially the same as those in Fig. 2: (i) the relationship of C_k to τ is parabolic-like for the original sequence, but it is horizontally linear for the shuffled sequence; (ii) when $1 \leq k \leq 4$, the parabolic-like of C_k is open downward, as opposed to the parabolic-like of C; (iii) the maximum correlation coefficients obtained at $k = 6$ for $a = 0, b = 0.02$ and $k = 7$ for $a = 0, b = 0.1$, respectively, as shown by the underlined numbers in Table 3.

Fig. 4 shows the influence of τ on C_k for the last case of $a \neq 0, b \neq 0$ with k from 1 to 12. More interestingly, the main results in Fig. 4 are essentially the same as those in Figs. 2 and 3: (i) the relationship of C_k to τ is parabolic-like for the original sequence, but it is horizontally linear for the shuffled sequence; (ii) when $1 \leq k \leq 4$, the parabolic-like of C_k is open upward, as opposed to the parabolic-like of C; (iii) the maximum correlation coefficients obtained at $k = 6$ for $a = b = 0.02$, $k = 6$ or $k = 7$ for $a = b = 0.1$ and $k = 7$ for $a = b = 0.5$, respectively, as depicted by the underlined numbers in Table 4.

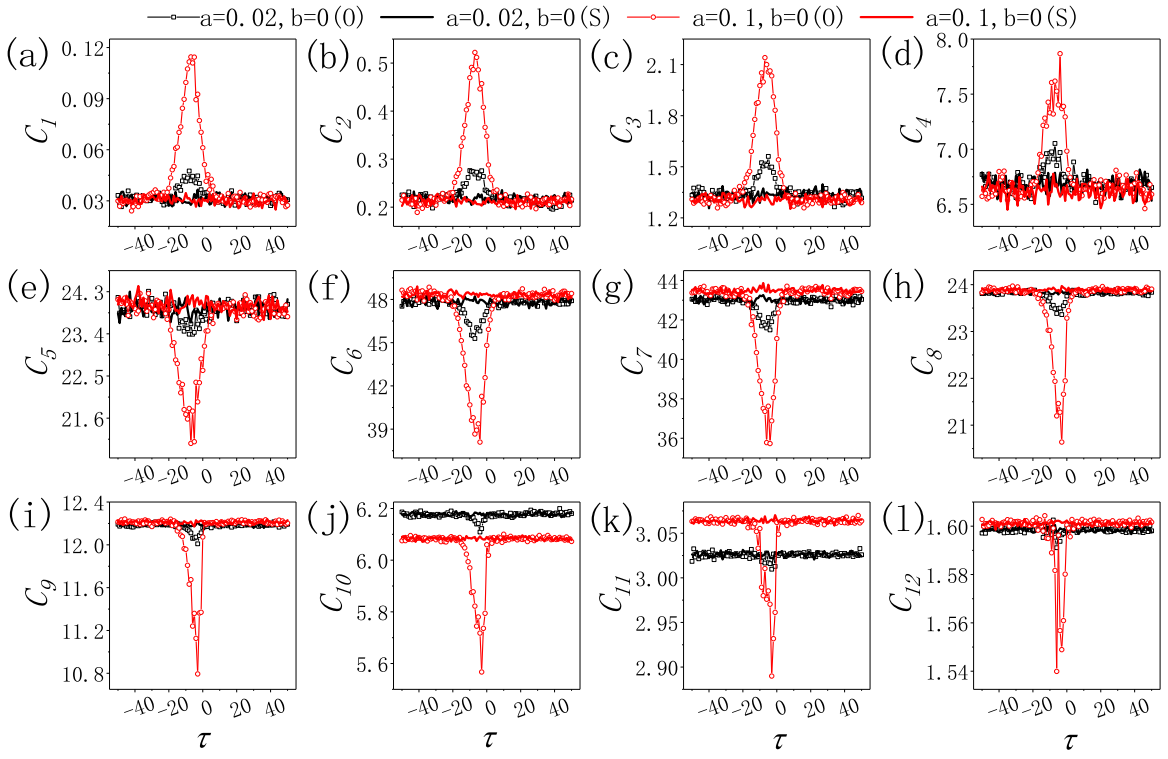

Fig. 2. The local complexity of the Logistic model (Eq. 2.1.1) for $a \neq 0, b = 0$. (a) to (l) are local complexity versus τ for the scale number k from 1 to 12, respectively. Here, black and red curves represent coupled parameters $a = 0.02, b = 0$ and $a = 0.1, b = 0$, respectively. Hollow shapes and solid lines represent original sequences (namely, "O") and shuffled sequences (namely, "S"), respectively.

Table 3
The correlation coefficient r between total complexity and local complexity of the Logistic model (Eq. 2.1.1) $a = 0, b \neq 0$.

The scale number k	$a = 0,\ b = 0.02$	$a = 0,\ b = 0.1$
	The correlation coefficient	The correlation coefficient
1	−0.89	−0.92
2	−0.93	−0.94
3	−0.96	−0.96
4	−0.92	−0.95
5	0.85	0.46
6	0.98	0.98
7	0.96	0.99
8	0.91	0.98
9	0.92	0.93
10	0.79	0.85
11	0.75	0.77
12	0.38	0.35

Hence, the results obtained in all three cases (i.e., $a \neq 0, b = 0$, $a = 0, b \neq 0$ and $a \neq 0, b \neq 0$) are consistent. To sum up, for Eq. 2.1.1, local complexity C_k contributes differently to total complexity C, with the medium scale contributing the most.

3.2. The Rössler–Lorenz model of unidirectional coupling

3.2.1. The influence of delay time on the total complexity

In this part, an application of the BS-MSC method to time series data generated by the Rössler–Lorenz model(Eq. (4)) is discussed. According to Eq. (4), the variable Y does not cause the variable V; however, whether the variable V causes the variable Y depends on the value of the coupling coefficient G. Three cases of values of G are examined, including $G = 0$, $G = 1$ and $G = 2$ in Fig. 5(a). The result illustrates that: (i) when $G = 0$, the relationship of C to τ is horizontally linear for both the original sequence and the shuffled sequence, i.e., the influence of τ on C is negligible. In other words, the variable V not causes the variable Y; (ii) when $G = 1$ or $G = 2$ (i.e., $G \neq 0$), the relationship of C to τ is parabolic-like with upward opening for the original sequence, but it is

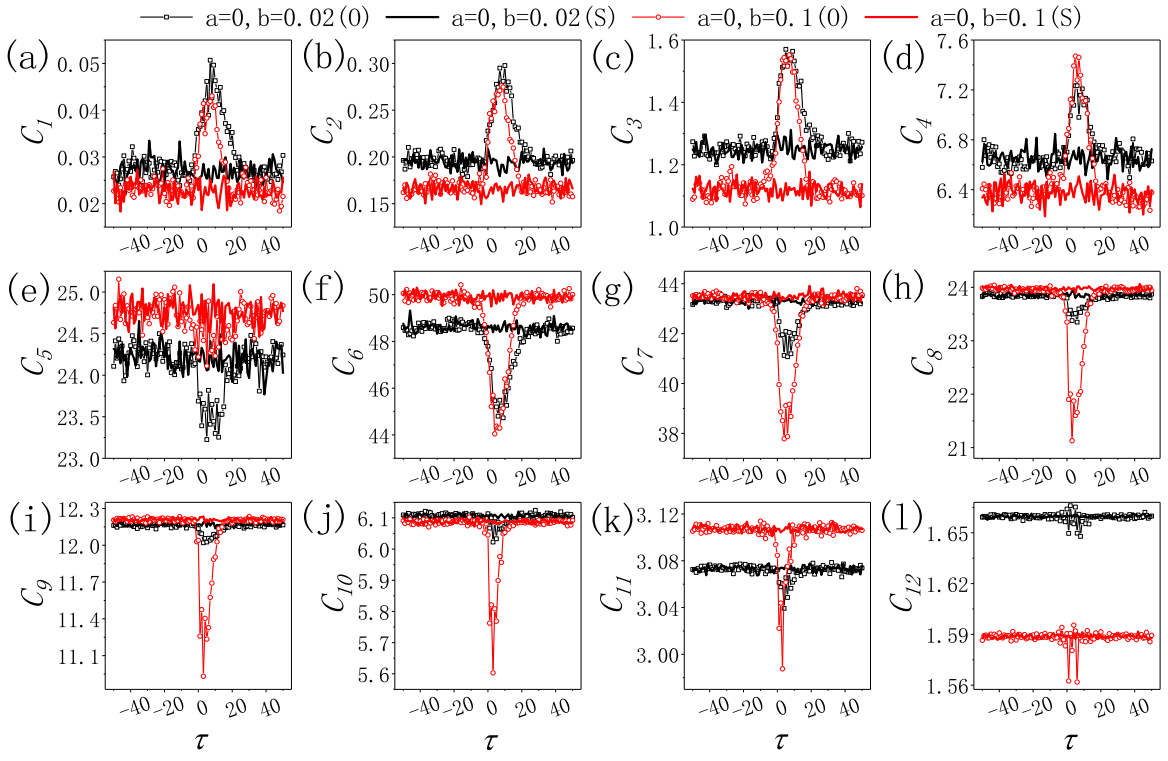

Fig. 3. The local complexity of the Logistic model (Eq. 2.1.1) for $a = 0, b \neq 0$. (a) to (l) are local complexity versus τ for the scale number k from 1 to 12, respectively. Here, black and red curves represent coupled parameters $a = 0, b = 0.02$ and $a = 0, b = 0.1$, respectively. Hollow shapes and solid lines represent original sequences (namely, "O") and shuffled sequences (namely, "S"), respectively.

Table 4
The correlation coefficient r between total complexity and local complexity of the Logistic model (Eq. 2.1.1) for $a \neq 0, b \neq 0$.

The scale number k	$a = b = 0.02$	$a = b = 0.1$	$a = b = 0.5$
	The correlation coefficient	The correlation coefficient	The correlation coefficient
1	−0.85	−0.96	−0.95
2	−0.91	−0.97	−0.97
3	−0.94	−0.98	−0.99
4	−0.91	−0.97	−0.96
5	0.68	0.53	0.90
6	0.99	0.99	0.99
7	0.96	0.99	1
8	0.95	0.97	0.98
9	0.94	0.94	0.96
10	0.84	0.95	0.93
11	0.78	0.93	0.89
12	0.43	0.76	0.75

horizontally linear for the shuffled sequence; (iii) when $G = 1$ or $G = 2$, a trough (here refers to the minimum value) exists for the original sequence, which value is $C_{min} = 149.34$ and $C_{min} = 135.38$ when $\tau = 6$ and $\tau = 2$ for $G = 1$ and $G = 2$, respectively. The trough values are much smaller than the complexity C for the shuffled sequence. Therefore, the variable V causes the variable Y because of $\tau > 0$; (iv) C_{min} for $G = 2$ is smaller than C_{min} for $G = 1$, i.e., the causality strength for $G = 2$ is larger than the causality strength for $G = 1$.

Compared to Eq. 2.1.1, the parabolic-like in Eq. (4) is not horizontally linear when τ is relatively large, i.e., there are several less obvious troughs beside the minimum trough C_{min}. Considering that the values of τ corresponding to these troughs are relatively large, we conjecture that the existence of these troughs does not arise from causality but rather from the period of the sequence itself. Next, verification is performed by calculating the periods of V and Y.

Many methods are available to analyze periodic components of signals, including power spectrum analysis [32–34], wavelet analysis [35], and empirical mode decomposition [36], of which power spectrum analysis is widely accepted as the most mature method. Fig. 6 shows the calculated period from the power spectrum analysis method. For V, no matter whether $G = 1$ (see Fig. 6(a)) or $G = 2$ (see Fig. 6(c)), two periods are observed: $p_{v_1} = 97.52$ and $p_{v_2} = 186.18$. It is due to the fact that V is independent of G. For

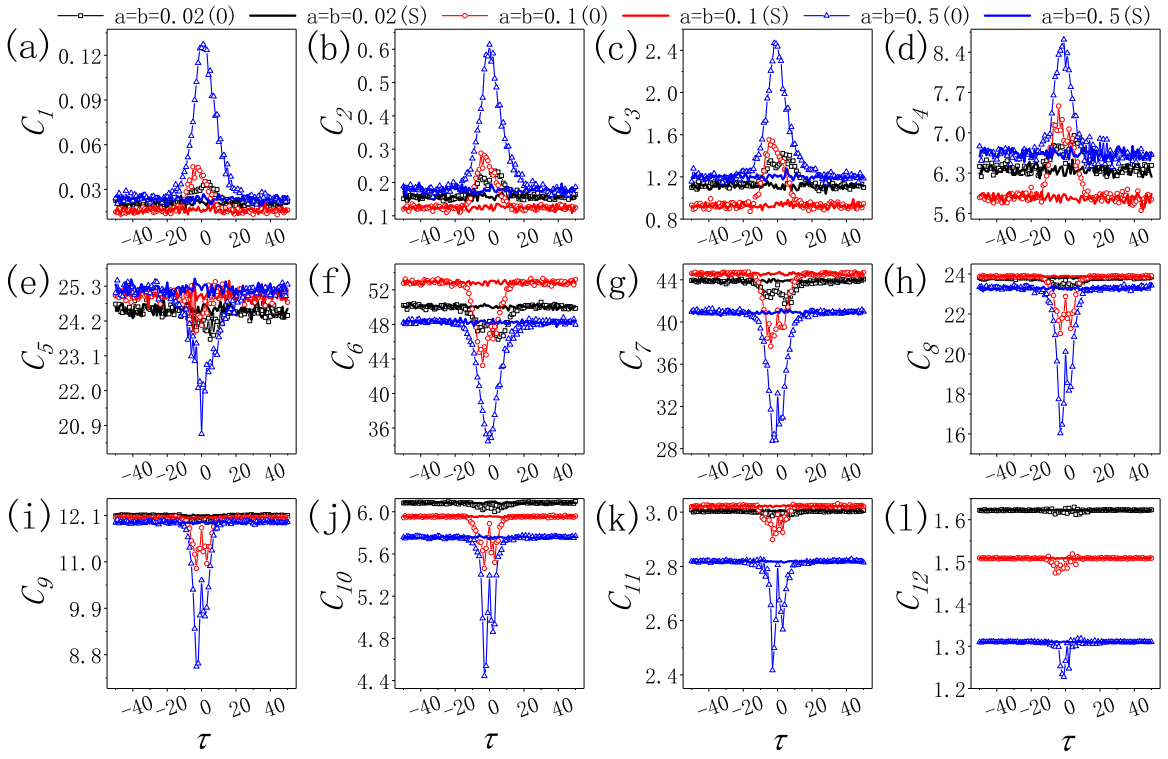

Fig. 4. The local complexity of the Logistic model (Eq. 2.1.1) for $a \neq 0, b \neq 0$. (a) to (l) are local complexity versus τ for the scale number k from 1 to 12, respectively. Here, black, red and blue curves represent coupled parameters $a = b = 0.02$, $a = b = 0.1$ and $a = b = 0.5$, respectively. Hollow shapes and solid lines represent original sequences (namely, "O") and shuffled sequences (namely, "S"), respectively.

Y, when $G = 1$, two periods are observed: $p_{y_1} = 4096$ (which is too large and should be discarded) and $p_{y_2} = 80.31$ (see Fig. 6(b)); when $G = 2$, two periods are observed: $p_{y_1} = 273.07$ and $p_{y_2} = 50.57$ (see Fig. 6(d)). Additionally, the complexity C between V and V and the complexity C between Y and Y are presented in Fig. 5(b) and Fig. 5(c), respectively. Fig. 5(b) shows that no matter whether $G = 1$ or $G = 2$, the period of C is $p_v = 26 + |-26| = 52$ (which is about twice p_{v_1}). Fig. 5(c) demonstrates that the periods of C are $p_y = 21 + |-21| = 42$ (which is about half of p_{y_2}) for $G = 1$ and $p_y = 18 + |-18| = 36$ (which is the result of the joint action of p_{y_1} and p_{y_2}) for $G = 2$, respectively. Nevertheless, C between V and Y is specifically related to the period of both V and Y, which is why it occurs in multiple troughs in Fig. 5(a).

3.2.2. The influence of delay time on the local complexity

In this part, for Eq. (4), we examined the variation of C_k with τ in Fig. 7 and the correlation coefficient between C and C_k in Table 5. We observed a similar favorable result that: (i) when $G = 0$, the relationship of C_k to τ is horizontally linear at each k for both the original sequence and the shuffled sequence; (ii) when $G = 1$ or $G = 2$ (i.e., $G \neq 0$), the relationship of C_k to τ is parabolic-like for the original sequence, but it is horizontally linear for the shuffled sequence. Specifically, when $1 \leq k \leq 4$, the parabolic-like of C_k is open downward, as opposed to a parabolic-like of C; when $k > 5$, the parabolic-like of C_k is open upward, just as that of C; (iii) the maximum correlation coefficients obtained at $k = 7$ for both $G = 1$ and $G = 2$, as shown by the underlined numbers in Table 5.

To sum up, for all three cases (i.e., $G = 0$, $G = 1$ and $G = 2$), the simulation results are consistent with the values of coupling coefficient G. Hence, for Eq. (4), local complexity C_k contributes differently to total complexity C, with the medium scale contributing the most.

3.3. Gait data

3.3.1. The influence of delay time on the total complexity

In order to demonstrate the applicability of the BS-MSC method to empirical data, we compared $\bar{C}$ of gait data for Parkinson's patients (Pt) with $\bar{C}$ of gait data for healthy subjects (Co), as shown in Fig. 8. Herein, $\bar{C}$ represents the mean value of C for each subjects within each specific subgroup (Group Pt consists of 29 subjects and Group Co consists of 18 subjects), where C for each subject is calculated by the sum of the eight sensor outputs of left foot and right foot. The result indicates that: (i) no matter whether Group Co or Group Pt, the relationship of $\bar{C}$ to τ is parabolic-like with upward opening for the original sequence, but it

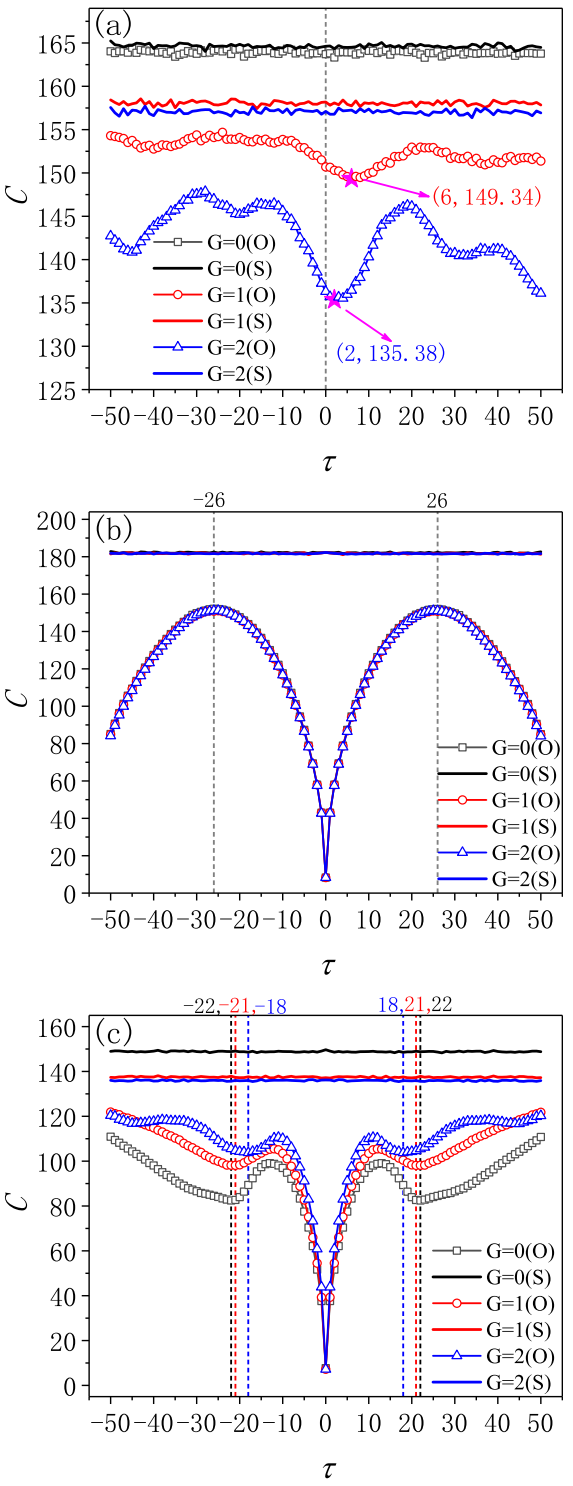

Fig. 5. The total complexity of the Rössler–Lorenz (Eq. (4)). (a) The total complexity between V and Y. (b) The total complexity between V and V. (c) The total complexity between Y and Y. Here, black, red and blue curves represent coupled parameters $G = 0$, $G = 1$, and $G = 2$, respectively. Hollow shapes and solid lines represent original sequences (namely, "O") and shuffled sequences (namely, "S"), respectively.

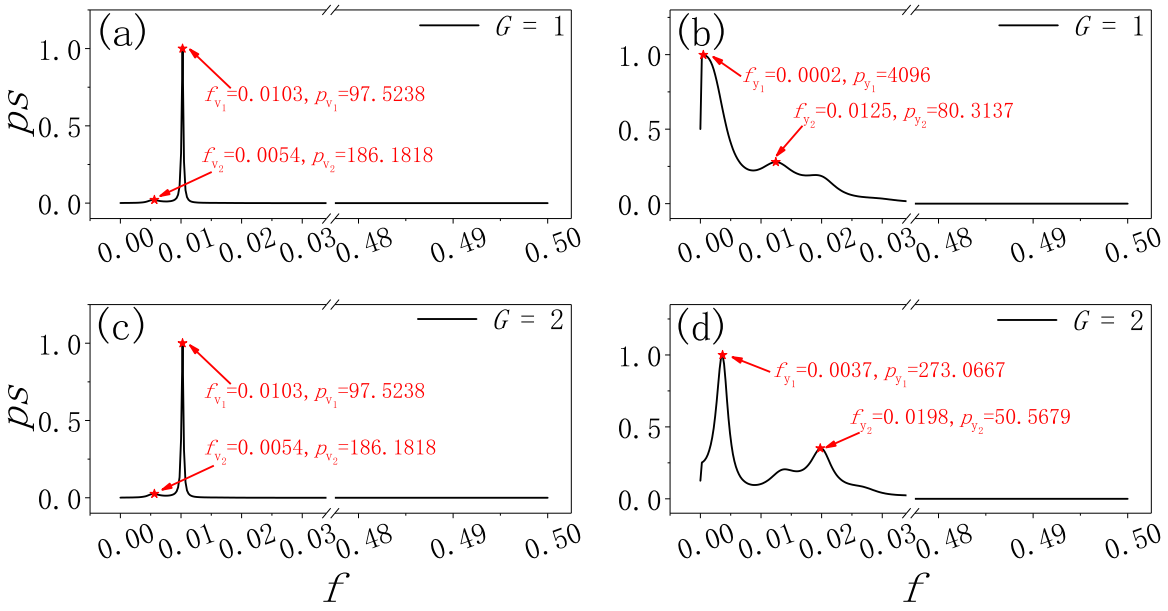

Fig. 6. The estimation of power spectrum. (a) The estimation of power spectrum for time series V when $G = 1$. (b) The estimation of power spectrum for time series Y when $G = 1$. (c) The estimation of power spectrum for time series V when $G = 2$. (d) The estimation of power spectrum for time series Y when $G = 2$.

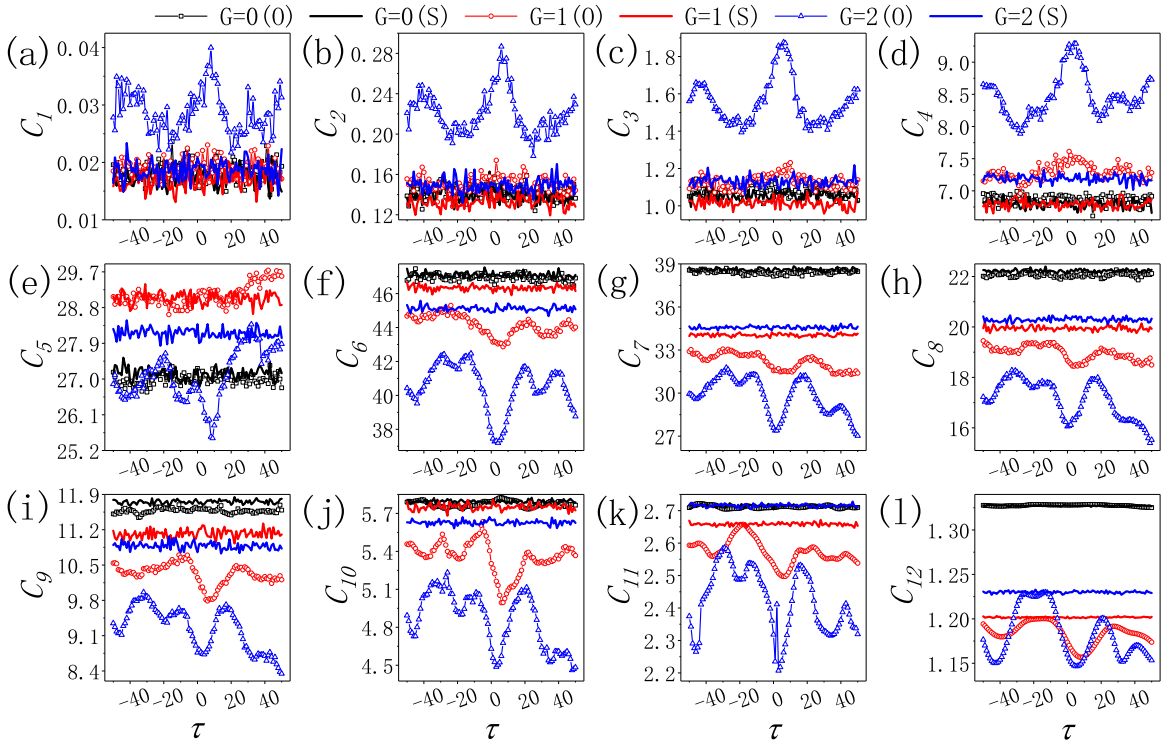

Fig. 7. The local complexity of the Rössler–Lorenz model (Eq. (4)). (a) to (l) are local complexity versus τ for the scale number k from 1 to 12, respectively. Here, black, red and blue curves represent coupled parameters $G = 0$, $G = 1$ and $G = 2$, respectively. Hollow shapes and solid lines represent original sequences (namely, "O") and shuffled sequences (namely, "S"), respectively.

Table 5
The correlation coefficient r between total complexity and local complexity of Rössler–Lorenz model (Eq. (4)) for $G = 1$ and $G = 2$.

The scale number k	$G = 1$	$G = 2$
	The correlation coefficient	The correlation coefficient
1	−0.21	−0.49
2	−0.31	−0.56
3	−0.53	−0.69
4	−0.60	−0.86
5	−0.44	0.04
6	0.90	0.90
7	0.91	0.96
8	0.90	0.90
9	0.84	0.87
10	0.75	0.89
11	0.84	0.88
12	0.83	0.79

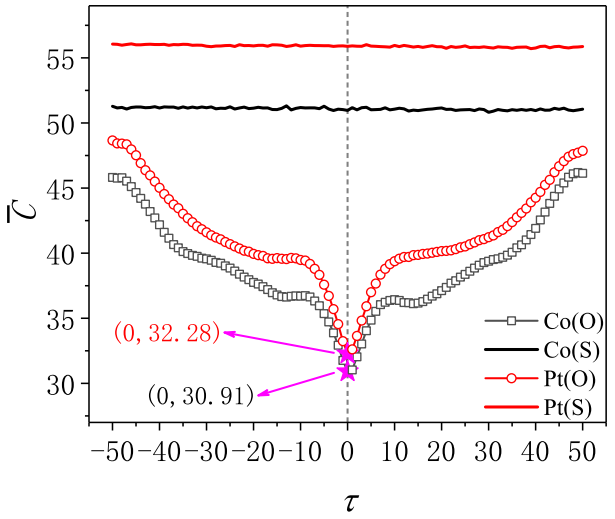

Fig. 8. The total complexity $\bar{C}$ of Gait data. Here, $\bar{C}$ represents the mean value of C for each subjects within each specific subgroup. Black and red curves represent healthy subjects (Co) and Parkinson's patients (Pt), respectively. Hollow shapes and solid lines represent original sequences (namely, "O") and shuffled sequences (namely, "S"), respectively.

is horizontally linear for the shuffled sequence. That is, there is a causality between the left foot and the right foot; (ii) a trough exists for the original sequence, which value is $\bar{C}_{min} = 30.91$ when $\tau = 0$ for Group Co and $\bar{C}_{min} = 32.28$ when $\tau = 0$ for Group Pt, respectively; (iii) the trough values of Group Co are much smaller than that of Group Pt, i.e., causalities between the left and right feet of Parkinson's disease patients are weaker than those of healthy individuals.

Research has found that gait fluctuations in healthy individuals are fractal properties or scale-invariant behaviors, but not in those with neurodegenerative conditions [29,30]. As found in our study, $\bar{C}$ for Group Pt is much larger than that for Group Co, which indicates that the gait fluctuations for Group Pt is more complicated than those for Group Co, i.e., deviate more from the fractal structure [15].

3.3.2. The influence of delay time on the local complexity

In this part, we compare the variation of $\bar{C}_k$ with τ for Group Pt with the variation of $\bar{C}_k$ with τ for Group Co. Just as $\bar{C}$, $\bar{C}_k$ represents the mean value of C_k for each subjects within each subgroup. Fig. 9 demonstrates that: (i) no matter whether Group Co or Group Pt, the relationship of $\bar{C}_k$ to τ is parabolic-like for the original sequence, but it is horizontally linear for the shuffled sequence; for the original sequence, (ii) when $1 \le k \le 3$, the parabolic-like of $\bar{C}_k$ is open downward, as opposed to the parabolic-like of $\bar{C}$; when $4 \le k \le 6$, although the parabolic-like of $\bar{C}_k$ is open upward, $\bar{C}_k$ for each τ is not always smaller than the value of $\bar{C}_k$ for the shuffled sequence; when $k = 7$, not only is the trend of $\bar{C}_k$ the same as $\bar{C}$, but the values of $\bar{C}_k$ are always smaller than the values of $\bar{C}_k$ for the shuffled sequence as well; (iii) the maximum correlation coefficients obtained at $k = 7$ for both Group Co and Group Pt, as shown by the underlined numbers in Table 6. Consequently, empirical data exhibit the same results as synthetic data, i.e., local complexity $\bar{C}_k$ contributes differently to total complexity $\bar{C}$, with the medium scale contributing the most.

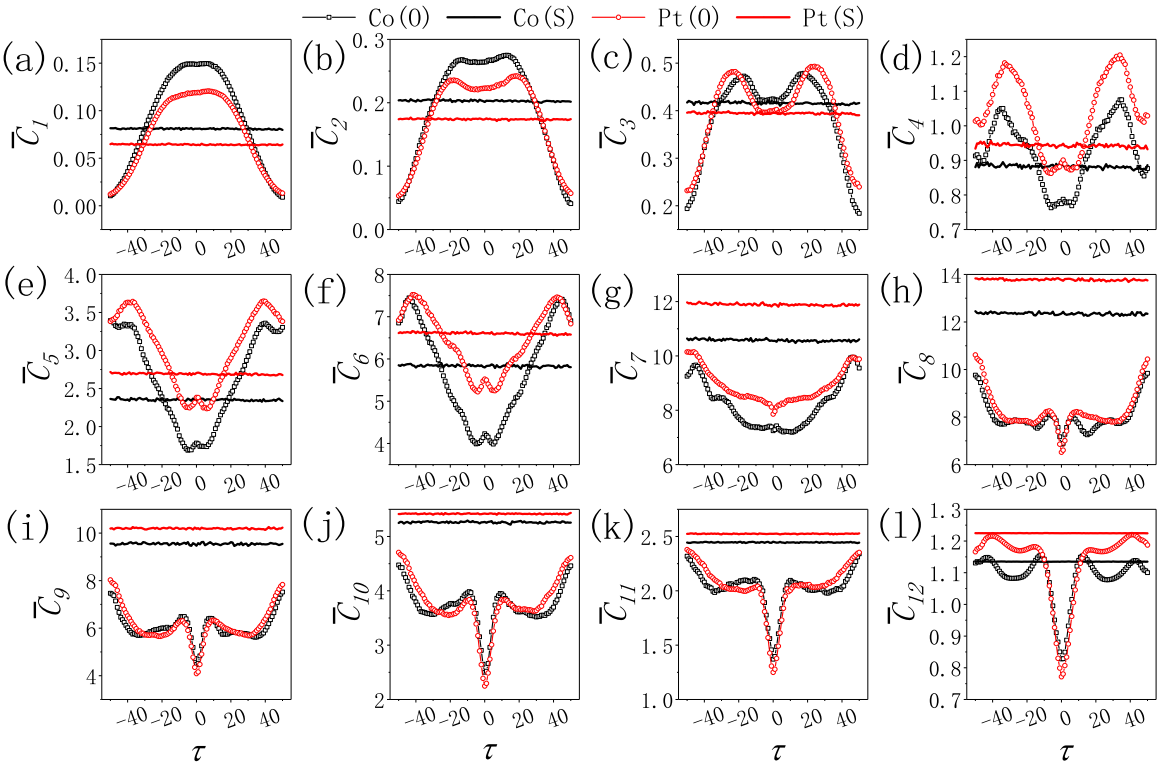

Fig. 9. The local complexity $\bar{C}_k$ of the gait data for Co and Pt. (a) to (l) are local complexity versus τ for the scale number k from 1 to 12, respectively. Here, $\bar{C}_k$ represents the mean value of C_k for each subjects within each subgroup. Black and red curves represent healthy subjects (Co) and Parkinson's patients (Pt), respectively. Hollow shapes and solid lines represent original sequences (namely, "O") and shuffled sequences (namely, "S"), respectively.

Table 6
The correlation coefficient r between mean total complexity and mean local complexity of gait data.

The scale number k	Co	Pt
	The correlation coefficient	The correlation coefficient
1	−0.95	−0.91
2	−0.94	−0.88
3	−0.85	−0.67
4	0.42	0.44
5	0.89	0.82
6	0.92	0.84
7	0.94	0.96
8	0.84	0.87
9	0.70	0.89
10	0.74	0.95
11	0.76	0.93
12	0.62	0.76

4. Conclusions and discussions

Several practical tools have emerged for measuring causality between variables in recent years [37–42]. These methods can distinguish the strength and direction of causality between variables, but they all share a limitation that they cannot capture the multi-scale characteristics of causality between variables. To achieve this goal, in this article we exploit recent advances in measuring the complexity of visual images to develop a multi-scale structural complexity method for characterizing causality between variables based solely on bivariate time series.

Firstly, we calculated the total complexity and local complexity of each scale for the synthetic sequences (including the time series obtained by the Logistic model of bidirectional coupling and the time series obtained by the Rössler–Lorenz model of unidirectional coupling). We found that: (i) the total complexity between variables can accurately portray the direction and strength of causality between them; (ii) the contribution of local complexity to the total complexity varies at different scales, with the medium scale

contribution being the largest. The results of our research are consistent with previous studies in that the scale-size should not be excessively large or small [14,15].

In addition, this conclusion holds not only for synthetic time series but also for empirical time series. By calculating the total or local complexity of Parkinson's patients and healthy subjects, we can accurately distinguish between them. The results shown that the causality strength between the left foot and right foot of Parkinson's patients is significantly weaker than that of healthy subjects.

To summarize, the BS-MSC method based solely on bivariate time series provides a novel perspective for the study of multi-scale causality between variables. This method, in addition to being a useful tool for studying causality between variables, can also provide a reference for future studies in choosing the appropriate scale, i.e., choosing the medium scale. The causal analysis conducted in this paper is focused on a bivariate relationship involving two variables. However, in the real-world context, the causality between two variables is inevitably influenced by a third variable, and potentially a more extensive set of variables. In other words, the causality between the two variables may be attributable to indirect causality resulting from mediation. Therefore, the mitigation of such indirect causality requires the implementation of conditional causal analysis. In future, it is worth studying the causality among three variables or the causality among multiple variables based on the BS-MSC method.

CRediT authorship contribution statement

Ping Wang: Conceptualization, Methodology, Data curation, Writing – original draft. **Changgui Gu:** Writing – review & editing, Funding acquisition. **Huijie Yang:** Supervision, Investigation. **Haiying Wang:** Visualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work is supported by the National Natural Science Foundation of China under Grant Nos. 12275179, 11875042 and 12150410309, and Natural Science Foundation of Shanghai, China (Grant No. 21ZR1443900).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.physa.2023.129398.

References

[1] G. Sugihara, R. May, H. Ye, C. Hsieh, E. Deyle, M. Fogarty, S. Munch, Detecting causality in complex ecosystems, Science 338 (2012) 496–500.
[2] S.F. Li, H. Zhang, D. Yuan, Investor attention and crude oil prices: Evidence from nonlinear Granger causality tests, Energy Econ. 8 (2019) 104494.
[3] T. Vyrost, S. Lyocsa, E. Baumohl, Granger causality stock market networks: Temporal proximity and preferential attachment, Physica A 427 (2015) 262–276.
[4] Q. Zhuo, Granger causal relations among Greater China stock markets: A nonlinear perspective, Appl. Financial Econ. 19 (2011) 1437–1450.
[5] K. Schiecke, F. Benninger, M. Feucht, Analysis of brain-heart couplings in epilepsy: Dealing with the highly complex structure of resulting interaction pattern, in: 28th European Signal Processing Conference, EUSIPCO, 2021, pp. 935–939.
[6] K. Schiecke, A. Schumann, F. Benninger, M. Feucht, K.J. Baer, P. Schlattmann, Brain-heart interactions considering complex physiological data: processing schemes for time-variant, frequency-dependent, topographical and statistical examination of directed interactions by convergent cross mapping, Physiol. Meas. 40 (2019) 114001.
[7] S.Y. Leng, H.F. Ma, W. Lin, L.N. Chen, Partial cross mapping eliminates indirect causal influences, Nature Commun. 11 (2020) 2632.
[8] C.W. Granger, Investigating causal relations by econometric models and cross-spectral methods, Econometrica 37 (1969) 424–438.
[9] S.X. Guo, C. Ladroue, J.F. Feng, Granger causality: Theory and applications, in: Frontiers in Computational and Systems Biology, Springer, New York, 2010, pp. 83–111.
[10] J.F. Geweke, Measures of conditional linear dependence and feedback between time series, J. Amer. Statist. Assoc. 79 (1984) 907–915.
[11] T. Schreiber, Measuring information transfer, Phys. Rev. Lett. 85 (2000) 461.
[12] R. Vicente, M. Wibral, M. Lindner, G. Pipa, Transfer entropy-a model-free measure of effective connectivity for the neurosciences, J. Comput. Neurosci. 30 (2011) 45–67.
[13] C. Cafaro, W.M. Lord, J. Sun, E.M. Bollt, Causation entropy from symbolic representations of dynamical systems, Chaos 25 (2015) 043106.
[14] X.T. Sun, W. Fang, X.Y. Gao, S.F. An, S.Y. Liu, T. Wu, Time-varying causality inference of different nickel markets based on the convergent cross mapping method, Resour. Policy 74 (2021) 102385.
[15] P. Wang, C.G. Gu, H.J. Yang, H.Y. Wang, Identify the characteristic in the evolution of the causality between the gold and dollar, Electron. Res. Arch. 30 (10) (2022) 3660–3678.
[16] C.W.J. Granger, J.L. Lin, Causality in the long run, Econom. Theory 11 (3) (1995) 530–536.
[17] E. Ghysels, J.B. Hill, K. Motegi, Testing for Granger causality with mixed frequency data, J. Econometrics 192 (1) (2016) 207–230.
[18] A.A. Bagrov, I.A. Iakovlev, A.A. Iliasov, M.I. Katsnelson, V.V. Mazurenko, Multiscale structural complexity of natural patterns, Proc. Natl. Acad. Sci. 117 (2020) 30241–30251.

[19] P. Wang, C.G. Gu, H.j. Yang, H.Y. Wang, The multi-scale structural complexity of urban morphology in China, Chaos Solitons Fractals 164 (2022) 112721.
[20] P. Wang, C.G. Gu, H.j. Yang, H.Y. Wang, J. Murdoch Mooreb, Characterizing systems by multi-scale structural complexity, Physica A 609 (2023) 128358.
[21] R.M. May, Simple mathematical models with very complicated dynamics, Nature 261 (1976) 459–467.
[22] A.L. Lloyd, The coupled logistic map: A simple model for effects of spatial heterogeneity on population dynamics, J. Theoret. Biol. 173 (1995) 217–230.
[23] E.N. Lorenz, Deterministic nonperiodic flow, J. Atmos. Sci. 20 (1963) 130–141.
[24] O.E. Rössler, An equation for continuous chaos, Phys. Lett. 57 (5) (1976) 397.
[25] O.E. Rössler, An equation for hyperchaos, Phys. Lett. A 71 (2–3) (1979) 155–157.
[26] H.-O. Peitgen, H. Jürgens, D. Saupe, Chaos and Fractals: New Frontiers of Science, Springer, New York, 2004, pp. 636–646.
[27] M. Le Van Quyen, J. Martinerie, C. Adam, F.J. Varela, Nonlinear analyses of interictal EEG map the brain interdependences in human focal epilepsy, Physica D 127 (1999) 250.
[28] A.L. Goldberger, PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals, Circulation 101 (23) (2000) e215–e220.
[29] J.M. Hausdorff, Altered fractal dynamics of gait: reduced stride-interval correlations with aging and Huntington's disease, J. Appl. Physiol. 82 (1997) 262–269.
[30] J.M. Hausdorff, A. Lertratanakul, M.E. Cudkowicz, A.L. Peterson, D. Kaliton, A.L. Goldberger, Dynamic markers of altered gait rhythm in amyotrophic lateral sclerosis, J. Appl. Physiol. 88 (6) (2000) 2045–2053.
[31] H. Ye, E.R. Deyle, L.J. Gilarranz, G. Sugihara, Distinguishing time-delayed causal interactions using convergent cross mapping, Sci. Rep. 5 (2015) 14750.
[32] J. Tang, X. Zhang, Optical variability periodicity analysis of BL Lac object S5 0716+714 based on bispectrum estimation, Acta Phys. Sin. 59 (10) (2010) 7516.
[33] G.L. Wang, Y.C. Liu, The method of power spectrum estimation based on measured pavement, in: International Conference on Electric Information and Control Engineering, 2011, pp. 2787–2790.
[34] Y. Yu, H.M. Zhao, A new method for noise power spectrum estimation, in: 4th IET International Conference on Wireless, Mobile and Multimedia Network, 2011, pp. 206–209.
[35] J. Morlet, G. Arens, E. Fourgeau, D. Giard, Wave propagation and sampling theory-part II: Sampling theory and complex waves, Geophysics 47 (1982) 222–236.
[36] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, The empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis, R. Soc. (454) (1998) 903–995.
[37] S.K. Stavroglou, A.A. Pantelous, H.E. Stanley, K.M. Zuev, Hidden interactions in financial markets, Proc. Natl. Acad. Sci. USA 116 (2019) 10646–10651.
[38] S.Y. Pranay, N. Nithin, Causal discovery using compression-complexity measures, J. Biomed. Inform. 117 (2021) 103724.
[39] S. Feizi, D. Marbach, M. Médard, M. Kellis, Network deconvolution as a general method to distinguish direct dependencies in networks, Nature Biotechnol. 31 (8) (2013).
[40] B. Barzel, A.L. Barabási, Network link prediction by global silencing of indirect correlations, Nature Biotechnol. 31 (8) (2013).
[41] X.T. Yu, G.J. Li, L.N. Chen, Prediction and early diagnosis of complex diseases by edge-network, Bioinformatics 30 (6) (2014) 852–859.
[42] J. Runge, P. Nowack, M. Kretschmer, S. Flaxman, D. Sejdinovic, Detecting and quantifying causal associations in large nonlinear time series datasets, Sci. Adv. 5 (2019) eaau4996.