Geometrical invariability of transformation between a time series and a complex network

Yi Zhao,[1,*] Tongfeng Weng,[1] and Shengkui Ye[2,3]

[1] *Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, China*

[2] *Mathematics and Physics Centre, Xi'an Jiaotong-Liverpool University, Suzhou, China*

[3] *Mathematical Institute, University of Oxford, Oxford, United Kingdom*

(Received 7 August 2013; revised manuscript received 14 March 2014; published 9 July 2014)

We present a dynamically equivalent transformation between time series and complex networks based on coarse geometry theory. In terms of quasi-isometric maps, we characterize how the underlying geometrical characters of complex systems are preserved during transformations. Fractal dimensions are shown to be the same for time series (or complex network) and its transformed counterpart. Results from the Rössler system, fractional Brownian motion, synthetic networks, and real networks support our findings. This work gives theoretical evidences for an equivalent transformation between time series and networks.

DOI: 10.1103/PhysRevE.90.012804 PACS number(s): 89.75.Hc, 05.45.Ac, 05.90.+m, 89.75.Fb

I. INTRODUCTION

Complex networks and time series are two generic ways to describe complex systems in the real world. Transformations from one to the other have attracted considerable attention [1–9]. Dynamical properties of time series are usually transformed into topological network structures via the similarities of the segmented cycle morphology [2], linear visibility conditions [3], amplitude difference of data points [4], or phase space reconstruction [6,7]. Since nodes of a network have no temporal sequence, the random walk strategy is usually adopted to obtain the temporal information of nodes [10,11] when the network is transformed into a time series. A deterministic method is proposed by Shimada *et al.* [12] to transform the ring lattices into sine data. However, whether the transformed network (or time series) can determine time-series dynamics (or network topological structure) is unknown to these transformation methods. It is not clear what dynamical characters are preserved during the given transformation and to what extent the given transformation is equivalent. Without an explicit transformation theory, it is hard to tell how the transformed network (or time series) characterizes time-series dynamics (or network topological structure). This is a fundamental problem but is less discussed.

In this paper we emphasize a dynamically equivalent transformation between time series and complex networks. Similarly, a discussion of equivalence between time series and their recurrence plots is presented in Ref. [13]. But this conclusion implies that the recurrence plot contains the temporal information of the time series, which is absent for the transformation from a network to a time series. Moreover, such equivalence is induced by topological homeomorphisms. However, topological homeomorphisms are weak conditions without considering the geometrical features of dynamical systems. For example, spatial distances, sizes, and shapes of dynamical systems are not preserved by homeomorphisms. A time series corresponds to a unique manifold in phase space, and thereby geometrical features play a key role in time-series dynamics. Meanwhile, geometrical concepts are widely employed to describe topological structures of complex networks [14–16]. We, therefore, propose geometrical invariability in a dynamically equivalent transformation. We theoretically show that there exists a quasi-isometric transformation where the geometrical features of a time series (or network) can be strictly preserved. Inspired by this finding, we adopt a correlation dimension [17] for describing a geometrical self-similarity of time-series dynamics and extend this concept to measure network dimension. We notice that the correlation dimension may have some limitations in characterization of a time series with multiscale features, as suggested in Ref. [18]. Gao *et al.* specifically proposed a scale-dependent measure to characterize various types of complex motions [19]. Here correlation dimension is used to explicitly exhibit preservation of geometrical characters and the existence of geometrical invariants during transformations. Additionally, we further show that for the fractional Brownian motion, there also exists a quasi-isometric transformation that ensures that the fractal dimension of the original stochastic time series is accurately captured by its network. We, therefore, provide theoretical evidences for a dynamically equivalent transformation, which is vital but absent in the previous studies.

II. TRANSFORMATION SCHEMATIC

Let $\{x_t\}_{t=1}^N$ be a scalar time series of N observations and ε be a threshold. We use the amplitude difference of data points as a transformation method, i.e., $a_{ij}=1$ if $|x_i-x_j|<\varepsilon$, otherwise $a_{ij}=0$. The matrix $A=\{a_{ij}\}_{i,j=1}^N$ is the adjacency matrix of the transformed network. Given this adjacency matrix, one can define a graphic distance matrix $D=\{d_{ij}\}_{i,j=1}^N$ of the same network. According to the classical multidimensional scaling (CMDS) [20], we get a square-distance matrix $S=\{d_{ij}^2\}_{i,j=1}^N$. Such a matrix S is then transformed to a centralizing gram matrix $G^c=-\frac{1}{2}HSH$, where $H=I-\frac{1}{N}E$, I is the $N\times N$ identity matrix, and E is the $N\times N$ matrix of ones (where all elements of this matrix equal to 1). Note that the defined graphic distance matrix D needs to be a Euclidean distance matrix such that G^c is a positive semidefinite matrix [21]. Let $\{\lambda_1,\lambda_2,\ldots,\lambda_N\}$ be the set of eigenvalues and $\boldsymbol{p}_i=\{p_{i1},p_{i2},\ldots,p_{iN}\}$ be the set of eigenvectors of G^c. Denote by h the number

*zhao.yi@hitsz.edu.cn

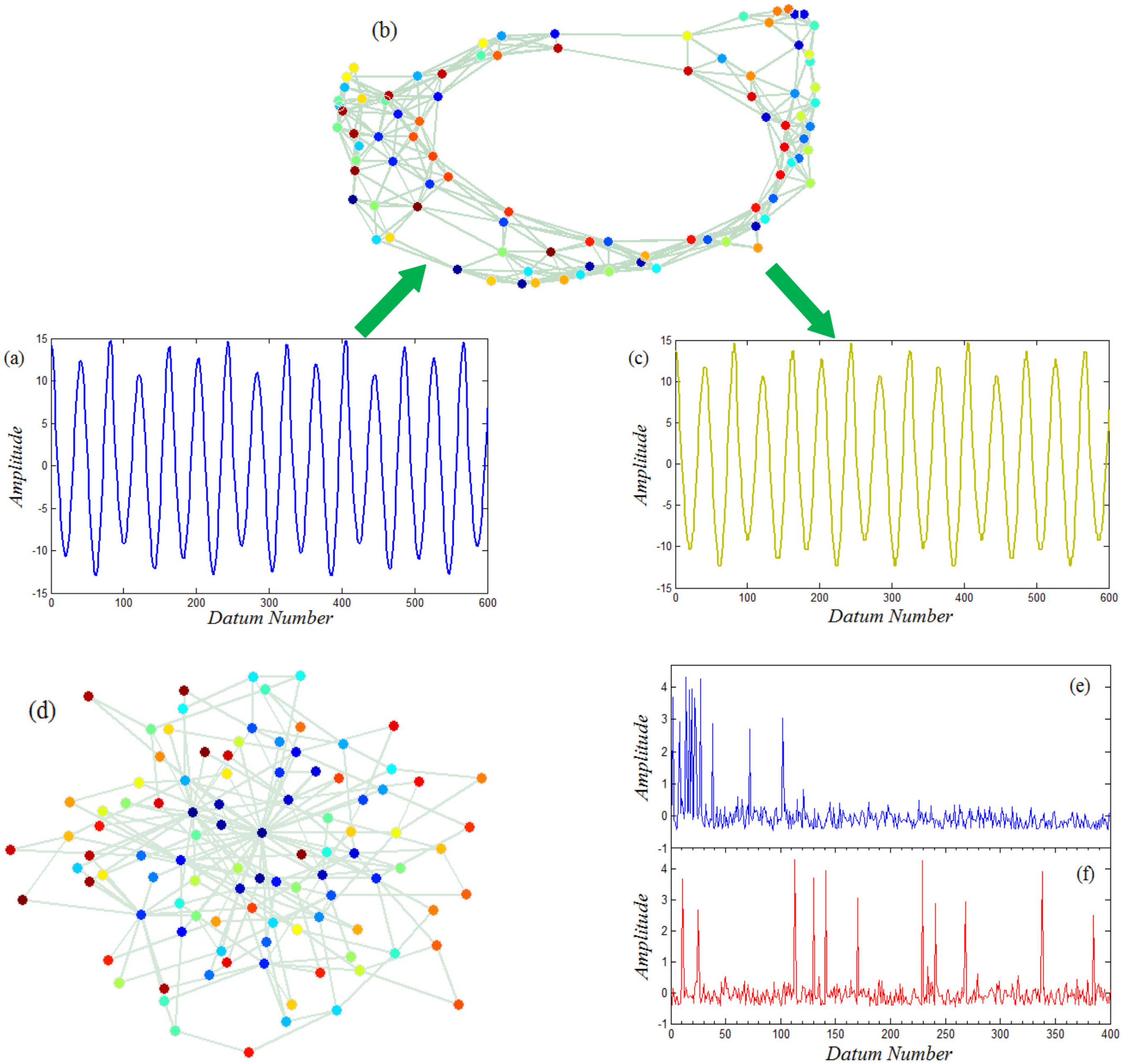

FIG. 1. (Color online) (a)–(c) Transformation from time series to its network. The x-component data of a Rössler system (a) in chaotic regime $[\dot{x} = -(y+z), \dot{y} = x + 0.2y, \dot{z} = 0.2 + 5.7(x - c)]$ with the iterative step size 0.2 is transformed into a network (b), provided that ε is small enough to ensure that the constructed network is a connected graph without isolated nodes or subgraphs. The time series (c) is then reproduced from the previous network. (d)–(f) Transformation from a network to its time series. A simple scale-free network (d) is transformed into two time series under two temporal strategies [i.e., the node sequence is determined by a random walk (e) or is assigned randomly (f)]. The latter is equivalent to randomly shuffle the former.

of nonzero eigenvalues. We perform spectral decomposition of the matrix $G^c = VV'$, where $V = (\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_h) = diag(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_h})(\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_h)$. The matrix V correspondingly forms a data set $\boldsymbol{Y} = (\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_N)' \in R^h$ and $\boldsymbol{y}_i = (y_{i1}, y_{i2}, \ldots, y_{ih})$. The matrix V is an exact configuration of the previous distance matrix D, i.e., the Euclidean norm of $\boldsymbol{y}_i$ and $\boldsymbol{y}_j$ equals d_{ij} [20]. In this sense, each network node is coordinated by a corresponding vector of Y, and finally all nodes form trajectories in a h-dimensional space. In practice, we usually choose a low-dimensional optimal configuration matrix in order to get a low-dimensional geometrical representation of the given network. In this case, the Euclidean distance of this new matrix is an approximation of D. Figure 1 depicts a schematic diagram of transformations between a time series and a network. Here we take the vector corresponding to the maximal eigenvalue as the reproduced time series.

It is interesting that in the transformation cycle (the top panel of Fig. 1) the original time series and the reproduced time series have obvious similarity. Their cross-correlation coefficient is almost 1. For transformation of a network into time series (the bottom panel of Fig. 1) using a random walk on the network, the transformed time series exhibits temporal correlation closely related to the network structure. However, we cannot reproduce exactly the same network as the original network from the transformed time series since this data is generated in a stochastic way. Using the quasi-isometric principle [22], we elaborate why the underlying dynamics of both the original and the reproduced time series are equivalent and how the time series transformed from a network preserves the network topological relationship.

III. THE QUASI-ISOMETRIC THEOREM OF TIME SERIES AND NETWORK

In this section we present mathematical evidences that two time series corresponding to the same network could be quasi-isometric. Let M and N be two closed subspaces of R^n for some positive integer $n < \infty$. Fix positive real values ε_1 and ε_2. Let φ be a surjection from the subspace M to the subspace N with the property that two points are ε_1-close in M, i.e., $d(x,y) < \varepsilon_1$, if and only if their images are ε_2-close in N, i.e., $d(\varphi(x),\varphi(y)) < \varepsilon_2$. Here d is the Euclidean metric. We also assume that the subspaces M and N satisfy the separation condition as follows. For each $x \neq y \in M$, there exists an element $z \in M$ such that $d(x,z) > \varepsilon_1$ and $d(y,z) < \varepsilon_1$. Similarly, whenever $u \neq v \in N$, there exists an element $w \in M$ such that $d(u,w) > \varepsilon_2$ and $d(v,w) < \varepsilon_2$.

For any two points $x,y \in M$, let d' denote the distance between x and y along M. Precisely, $d'(x,y)$ is defined as the infimum of lengths of all paths connecting x and y in M. Similarly, we define a similar metric d' on N. To eliminate confusion, we write d instead of d' for short. Our target is to show that φ is a quasi-isometry, i.e., $\frac{\varepsilon_2}{\varepsilon_1}d(x,y) - \varepsilon_2 < d(\varphi(x),\varphi(y)) < \frac{\varepsilon_2}{\varepsilon_1}d(x,y) + \varepsilon_2$ for any two elements $x,y \in M$.

First, it is easy to justify that φ is injective. Since for two arbitrary distinct points $\varphi(x)$ and $\varphi(y)$ in N, based on the separation condition, there is an element $\varphi(z) \in N$ with $d(\varphi(x),\varphi(z)) > \varepsilon_2$ and $d(\varphi(y),\varphi(z)) < \varepsilon_2$. In return we have $d(x,z) > \varepsilon_1$ and $d(y,z) < \varepsilon_1$ in the subspace M. So φ is injective.

Now we prove that φ is a quasi-isometry. For two arbitrary points x and y in the subspace M, there is an integer k such that

$$k\varepsilon_1 \leqslant d(x,y) < (k+1)\varepsilon_1. \quad (1)$$

We can find a path P connecting x and y whose length is very close to $d(x,y)$. Choose a partition $x = x_0,x_1,\ldots,x_{n+1} = y$ of P such that $d(x_i,x_{i+1}) < \varepsilon_1$ for each $i = 0,1,\ldots,n$. Since $d(x,y) < \varepsilon_1$ if and only if $d(\varphi(x),\varphi(y)) < \varepsilon_2$, we get $\varphi(B(z,\varepsilon_1)) = B'(\varphi(z),\varepsilon_2)$, where $B(z,\varepsilon_1)$ is an open ball in M with center z and radius ε_1. Therefore, by triangle inequalities we have

$$d(\varphi(x),\varphi(y)) \leqslant \sum_{i=0}^{n} d(\varphi(x_i),\varphi(x_{i+1})) < (k+1)\varepsilon_2. \quad (2)$$

It can be rewritten as

$$d(\varphi(x),\varphi(y)) < k\varepsilon_1\frac{\varepsilon_2}{\varepsilon_1} + \varepsilon_2 \leqslant \frac{\varepsilon_2}{\varepsilon_1}d(x,y) + \varepsilon_2. \quad (3)$$

By symmetric relationship between subspaces M and N, we also have

$$d(x,y) < \frac{\varepsilon_1}{\varepsilon_2}d(\varphi(x),\varphi(y)) + \varepsilon_1. \quad (4)$$

By combining Eqs. (3) and (4) together, we obtain the following relations:

$$\frac{\varepsilon_2}{\varepsilon_1}d(x,y) - \varepsilon_2 < d(\varphi(x),\varphi(y)) < \frac{\varepsilon_2}{\varepsilon_1}d(x,y) + \varepsilon_2. \quad (5)$$

Thus φ is a quasi-isometry.

Now back to our problem, we assume that there are two scalar time series $M = \{x_i\}_{i=1}^N$ and $M' = \{x_i'\}_{i=1}^N$ corresponding to the same network. In other words, for some positive real numbers ε_1 and ε_2, they produce the same adjacency matrix of a transformed network, i.e., $|x_i - x_j| < \varepsilon_1$ if and only if the corresponding points satisfy the inequality $|x_i' - x_j'| < \varepsilon_2$. As discussed previously, we assume that M (and M') satisfies the separation property that for each $x_i \neq x_j$ in M, there exists an element $x_k \in M$ such that $|x_i - x_k| < \varepsilon_1$ and $|x_k - x_j| > \varepsilon_1$. For any two points $x,y \in M$, we find a path P connecting x and y in M, whose length is close to the distance $|x-y|$. Divide the path P into small pieces $x = x_0 \to x_1 \to x_2 \to \cdots \to x_n = y$ such that for each i we have $|x_i - x_{i+1}| < \varepsilon_1$. Since M and M' induce the same adjacency matrix, the corresponding points x', y' in M' can be connected by a path $x' = x_0' \to x_1' \to x_2' \to \cdots \to x_n' = y'$ such that $|x_i' - x_{i+1}'| < \varepsilon_2$. Therefore, the distance $|x' - y'| < n\varepsilon_2$, where the right-hand side is close to $\frac{\varepsilon_2}{\varepsilon_1}|x-y|$. Actually, we can prove that

$$\frac{\varepsilon_2}{\varepsilon_1}|x-y| - \varepsilon_2 < |x'-y'| < \frac{\varepsilon_2}{\varepsilon_1}|x-y| + \varepsilon_2. \quad (6)$$

These inequalities indicate that if M and M' induce the same network, the distances between corresponding points are almost the same as well when the small numbers ε_1 and ε_2 are close to each other. We say that there is a quasi-isometric map between M and M'. Based on that, we can observe a sort of geometrical invariability during transformations as follows.

IV. NUMERICAL SIMULATIONS

To illustrate a dynamically equivalent transformation between a time series and a complex network, we take a chaotic Rössler time series of 10 000 data points with the step size of 0.2 as an example. Following the above procedure, we get the adjacency matrix $A = \{a_{ij}\}$ of the transformed network. The threshold for this transformation is denoted by ε_1. For any adjacent nodes i and j, let $G_i = \{j : a_{ij} = 1\}$ denote the set of nodes having links to node i. The weight w_{ij} is defined as $1 - \frac{|G_i \cap G_j|}{|G_i \cup G_j|}$, where $|G_i|$ denotes the cardinality of the set G_i [9]. We get an existing link having weights $W = \{w_{ij}\}$ and a graphic distance matrix $D = \{d_{ij}\}$, where d_{ij} is the shortest distance between nodes i and j based on these new link weights. The CMDS is equivalent to the principal components analysis when handling the data set [20]. We use the eigenvector corresponding to the maximal eigenvalue to produce a time series since the maximal eigenvalue contributes 90% of the sum of all nonzero eigenvalues.

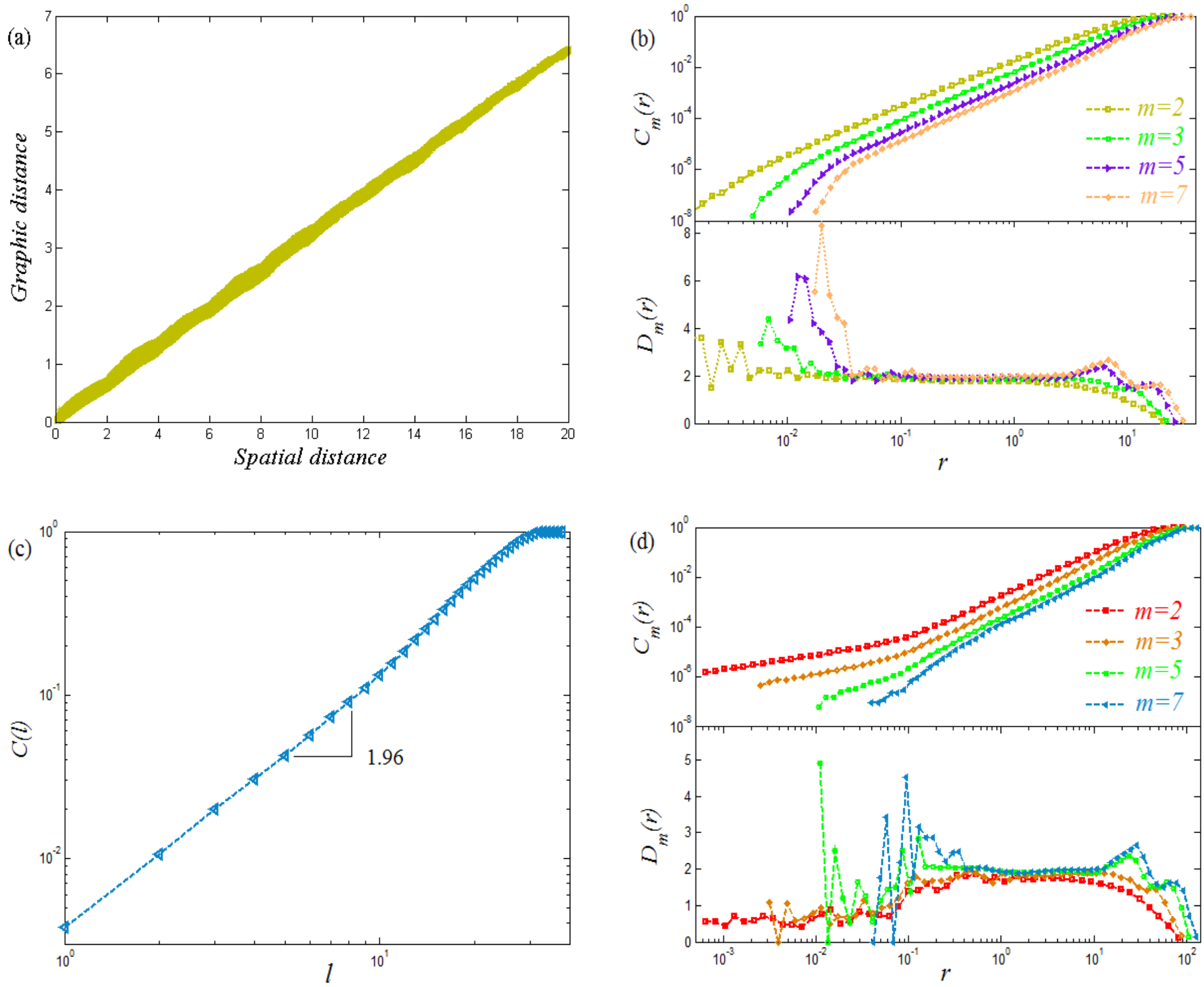

FIG. 2. (Color online) (a) The Euclidean distance between the Rössler data points in phase space versus graphic distances of the corresponding network nodes. (b) (Top panel) Log-log plot of correlation sum $C_m(r)$ as a function of the distance scalar r. Different curves are obtained with embedding dimensions of $m = 2$ (uppermost curve) to $m = 7$ (lowest curve). (Bottom panel) Local slopes of the correlation sums shown in the upper panel. The correlation dimension is an estimate of $D = 2 \pm 0.05$. (c) Correlation sum $C(l)$ for the transformed network is plotted versus the shortest path length l. The network correlation dimension is estimated to 1.96. (d) (Top panel) Log-log plot of the correlation sum $C_m(r)$ versus the distance scalar r for the reconstructed data. (Bottom panel) Local slopes of the correlation sums shown in the upper panel. The correlation dimension is an estimate of $D = 1.98 \pm 0.05$.

From Fig. 2, we observe that there is a strictly linear relationship between the Euclidean distance of data points in phase space and the graphic distance of corresponding nodes. This suggests that the network structure contains the underlying geometrical features of the original time-series dynamics. Moreover, the correlation dimension of the Rössler system is consistent with that of the transformed network, as well as that of the reproduced time series. We calculate a correlation dimension of time series according to the Grassberger and Procaccia algorithm [17]. Note that calculation of correlation dimension is sensitive to the step sizes of the Rössler system. The large step size indicates coarse discretization of the Rössler system, and then there will be fewer neighbors within the given threshold. Statistical errors have an obvious effect in the small scaling levels, and thereby the middle scaling regime likely disappears.

For calculation of a network's correlation dimension, we randomly select one seed node and then count nodes whose shortest path lengths to this seed node are smaller than l. This procedure is repeatable by choosing each node as a seed node. The average number of nodes centered around the seed nodes is a function of l (i.e., the network fractal cluster dimension [23]). This calculation can be regarded as an unbiased estimation of network correlation dimension, in contrast to the calculation procedure in Refs. [14,24].

We can choose another threshold ε_2 such that the transformed network can be reconstructed from the reproduced time series. Notice that ε_2 is normally close to ε_1 and both are very

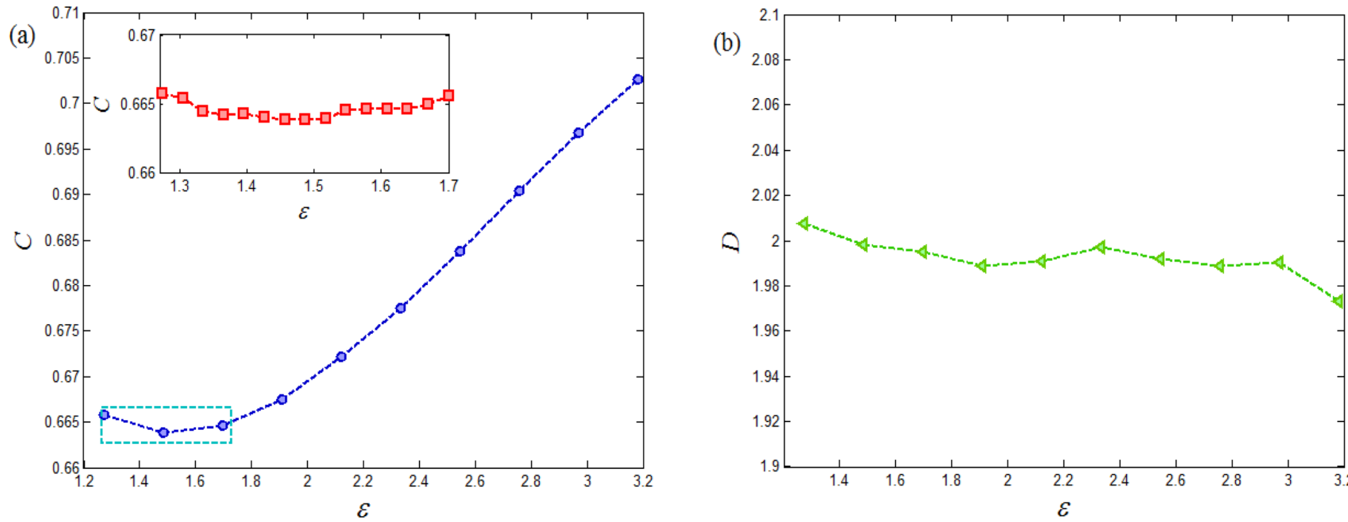

FIG. 3. (Color online) The effect of threshold value ε on network structure: (a) clustering coefficient C and (b) correlation dimension D. Inset: local view of the box area.

small. By the inequalities of quasi-isometry, we conclude that the original and reproduced time series are quasi-isometric. When ε_1 approximately equals ε_2 and both are out of the scaling regime, we further theoretically confirm that both the original and the reproduced time series corresponding to the same network structure shall have the same correlation dimension.

Recall that the correlation sum for any set of n points in an m-dimensional space $\{\vec{x}_i\}_{i=1}^{n}$ is defined as [17]

$$C(\varepsilon) = \lim_{n\to+\infty} \frac{g(\varepsilon)}{n^2}, \tag{7}$$

where $g(\varepsilon) = \sum_{i,j=1}^{n} \Theta(\varepsilon - |\vec{x}_i - \vec{x}_j|)$ represents the total number of pairs of points whose distance is less than ε, and $\Theta(\cdot)$ is the Heaviside function. In the limit of an infinite amount of data and for small ε, a scaling regime and a scaling for the correlation integral appear, i.e., $C(\varepsilon) \sim \varepsilon^D$. The exponent D is defined as the correlation dimension of time series.

Suppose that $f: M \to N$ is a quasi-isometric map of two sets of points, i.e., $K d_M(x,y) - c \leqslant d_N(f(x), f(y)) \leqslant K d_M(x,y) + c$ for some constants $K > 0$ and $c \geqslant 0$. We note that the variables in space M and N are denoted with the subscripts M and N, respectively. Without loss of generality, we assume that f is injective. Suppose that two points $x, y \in M$, satisfying $d_M(x,y) < \varepsilon$; then $d_N(f(x), f(y)) \leqslant K\varepsilon + c$. This proves that $g_M(\varepsilon) \leqslant g_N(K\varepsilon + c)$. Therefore, when ε and $K\varepsilon + c$ are small,

$$\begin{aligned} D_M &= \log_\varepsilon \varepsilon^{D_M} = \log_\varepsilon C(\varepsilon) = \lim_{n\to+\infty} \log_\varepsilon \frac{g_M(\varepsilon)}{n^2} \\ &\leqslant \lim_{n\to+\infty} \log_\varepsilon \frac{g_N(K\varepsilon + c)}{n^2} \\ &\leqslant \lim_{n\to+\infty} \log_\varepsilon \frac{g_N(K(\varepsilon + c/K))}{n^2} \\ &= \log_\varepsilon (K(\varepsilon + c/K))^{D_N} \\ &= D_N(\log_\varepsilon K + \log_\varepsilon(\varepsilon + c/K)). \end{aligned}$$

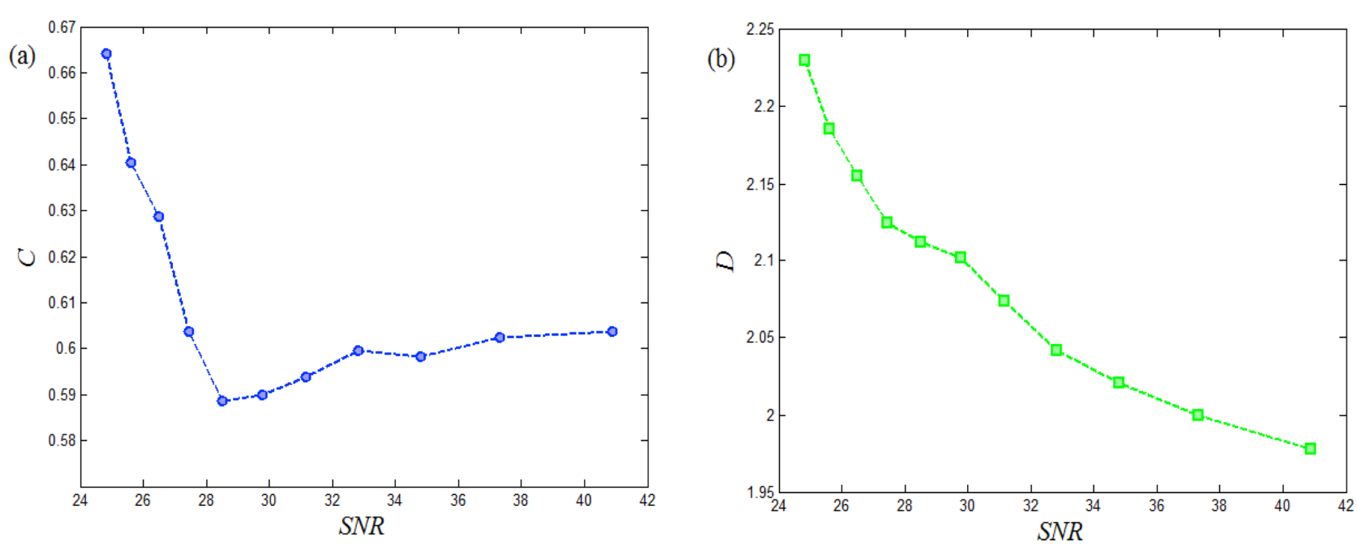

FIG. 4. (Color online) The effect of noise level on the network structure: (a) clustering coefficient C and (b) correlation dimension D.

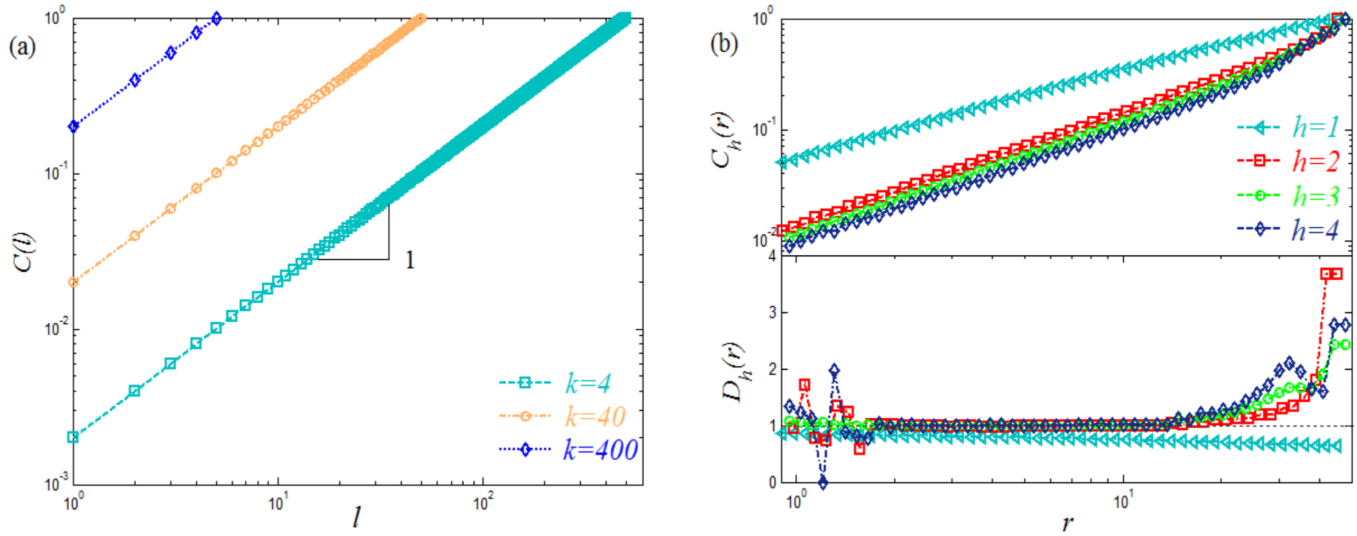

FIG. 5. (Color online) (a) Log-log plot of correlation sum $C(l)$ as a function of l for ring lattice networks with each node having k links. Correlation dimensions of these ring lattice networks equal 1. (b) (Top panel) Log-log plot of correlation sum $C_h(r)$ as a function of the distance scalar r for a series of multidimensional vectors reconstructed from a ring lattice network ($k = 40$) for different values of h. (Bottom panel) Local slopes of the correlation sums shown in the upper panel. The correlation dimension is an estimate of $D = 1 \pm 0.03$. The convergent value is marked by a horizontal dashed line.

With reference to the previous section, here $K = \varepsilon_2/\varepsilon_1$ and $c = \varepsilon_2$. If M and N are time series corresponding to the same network, assume that $\varepsilon_1 = \varepsilon_2$ are very small compared with ε. This is a reasonable assumption since in practice ε cannot be below a length scale of a few multiples of noise level (if the data is noisy) or discretization accuracy, and moreover, at small accessible scales, ε needs to ensure adequate neighbors [25]. The scaling regime is a finite intermediate interval, while ε_1 and ε_2 could be arbitrary small. With these assumptions, the right-hand side of the last equality is D_N. With the same argument, we prove that $D_N \leqslant D_M$. This shows that D_M and D_N are the same.

Technically, for a given time series, it is still a challenge to define the threshold value small enough. To address this, we investigate the effect of threshold value ε on the structure of transformed network. We take the previous chaotic Rössler time series as an example. To better describe the network structure, we adopt a network statistic, i.e., clustering coefficient C. As shown in Fig. 3, there is a toleration range for the minimal transformation threshold so that the network in terms of clustering coefficient is almost kept unchanged. Moreover, the correlation dimension of the transformed networks is consistent with that of the Rössler system. We also note that the larger threshold value results in the smaller scaling regime

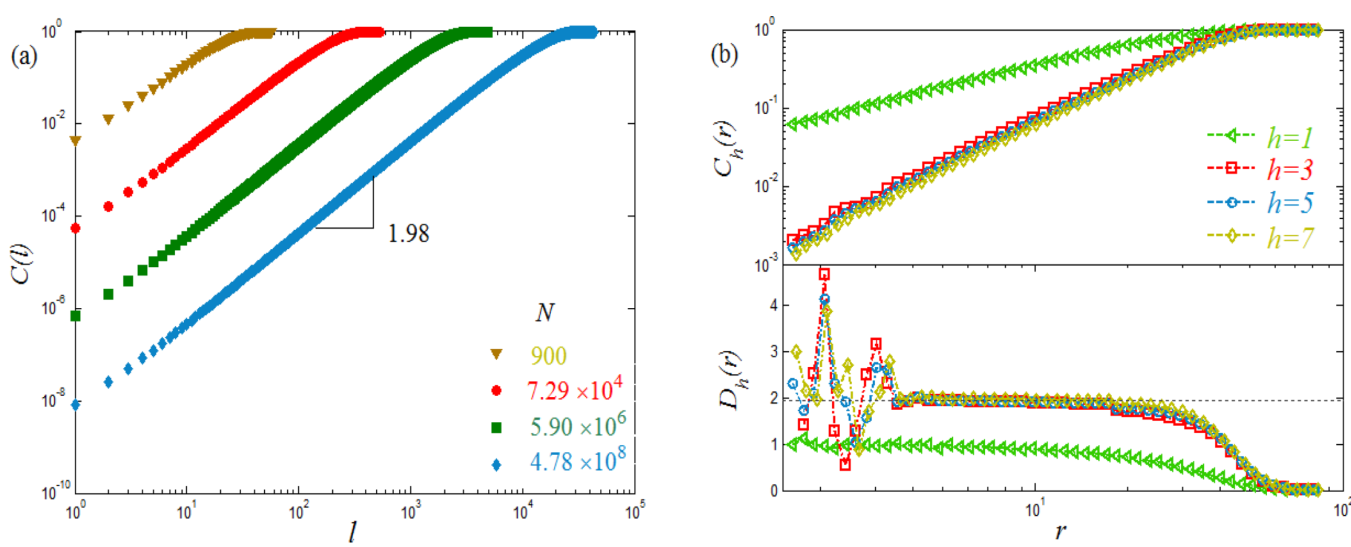

FIG. 6. (Color online) (a) Log-log plot of correlation sum $C(l)$ as a function of the shortest path length l for 2D lattice networks with N nodes. Correlation dimensions of these 2D lattice networks are estimated to 1.98. (b) (Top panel) Log-log plot of correlation sum $C_h(r)$ as a function of a distance scalar r for a series of h-dimensional vectors reconstructed from a 2D lattice network. (Bottom panel) Local slopes of the correlation sums shown in the upper panel. The correlation dimension is an estimate of $D = 1.98 \pm 0.03$. The convergent value is marked by a horizontal dashed line.

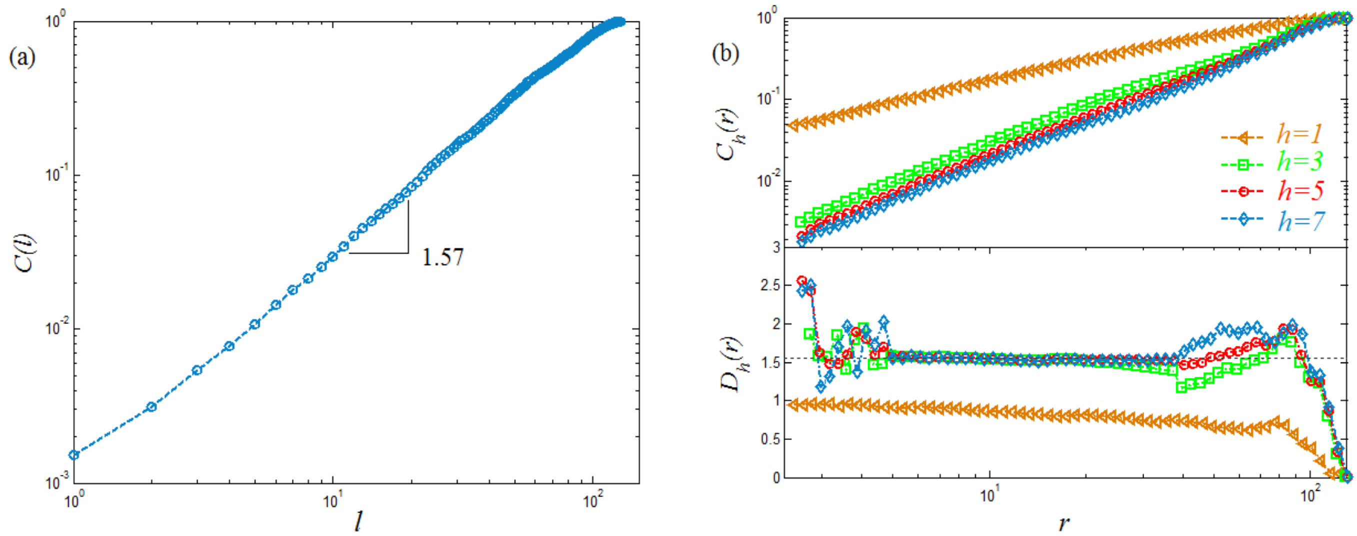

FIG. 7. (Color online) (a) Log-log plot of correlation sum $C(l)$ as a function of the shortest path length l for the Sierpinski gasket with the 8th generation. Correlation dimension of the Sierpinski gasket is estimated to 1.57. (b) (Top panel) Log-log plot of correlation sum $C_h(r)$ as a function of a distance scalar r for a series of h-dimensional vectors reconstructed from the Sierpinski gasket. (Bottom panel) Local slopes of the correlation sums shown in the upper panel. The correlation dimension is an estimate of $D = 1.57 \pm 0.03$. The convergent value is marked by a horizontal dashed line.

of the correlation sum. Hence, the slight variation of threshold value ε has little influence on network structure.

Moreover, we test the robustness of our transformation method against observational Gaussian noise given the same threshold as that in Fig. 2. Figure 4 shows the clustering coefficient versus noise level (signal-to-noise ratio, SNR) for the previous Rössler time series. It is shown that the clustering coefficient is approximately stable at SNR > 30 dB. Meanwhile, the correlation dimension of the corresponding transformed networks is approximately 2. However, along with more noise, the correlation dimension of the transformed networks deviates significantly from the correlation dimension of the chaotic Rössler system, suggesting that the quasi-isometric condition cannot be satisfied. Meanwhile, the clustering coefficient also indicates that the network structure changes significantly. In summary, the transformation theory shows its robustness in practical conditions where the threshold varies somewhat or the time series is contaminated with a weak noise.

We now study transformations from three typical synthetic networks with well-defined dimensions. The first one is a ring lattice network with each node having k links [26]. For a randomly chosen seed node, the number of nodes centered at the seed node with distance less than l is $l \times k$. When the distance increases by 1, the number of nodes centered at the seed node increases to $k(l+1)$. It is clear that the average number of nodes $C(l)$ shows a linear scaling with respect to the distance, i.e., $C(l) \propto l$. This implies that the dimension of a ring lattice network equals that of the straight line in space. Here we test a ring lattice network with 10,000 nodes. As expected, in Fig. 5(a) $C(l)$ shows a scaling, whose slope is the same as the theoretical value. By CMDS, this network is transformed into the multidimensional vectors. This process can be understood as reconstruction of a network in phase space with the multidimension denoted by h as an embedding dimension. We select the vectors corresponding to the maximal h eigenvalues to compute the spatial distance between nodes by the Euclidean norm. According to Fig. 5(b), correlation dimension of the multidimensional vectors is consistent with that of the ring network. This suggests that these vectors can accurately represent the geometrical features of this network.

A two-dimensional (2D) lattice network can be regarded as a discretion limit of a smooth metric plane [16]. Therefore, a correlation dimension of a 2D lattice network with finite but adequate nodes is almost two (i.e., the Hausdorff dimension of a plane). We show correlation sums of 2D lattice networks in Fig. 6(a). Correlation sums of a series of the multidimensional vectors from a 2D lattice network with 10,000 nodes corresponding to the maximal h eigenvalues are presented in Fig. 6(b). Correlation dimensions saturate

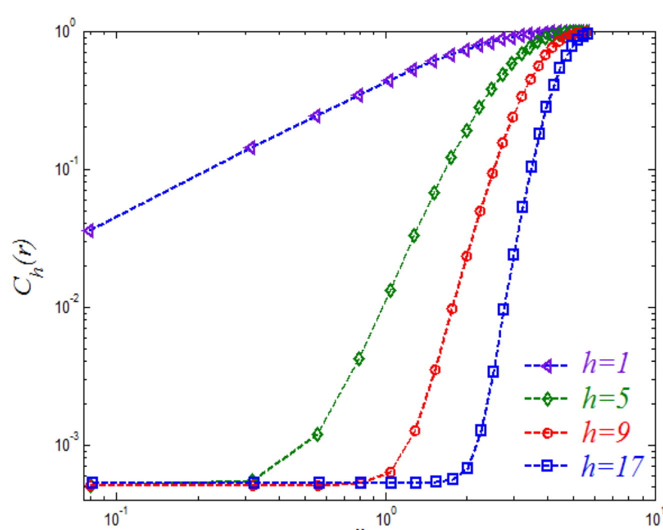

FIG. 8. (Color online) Log-log plot of correlation sum $C_h(r)$ as a function of the distance scalar r for a series of the h-dimensional vectors from a random network with 10,000 nodes. The scaling keeps increasing with larger h.

and approximately converge to 1.98 when h is more than 3. The maximal five eigenvalues dominate 90.3% of the sum of all nonzero eigenvalues. The space representation of a 2D lattice network described by these multidimensional vectors constitutes a flat surface, showing the same geometry as the given network.

We further validate another regular network (i.e., Sierpinski gasket) with a known fractal dimension. We generate a Sierpinski gasket composed of 3282 nodes and 6561 edges. From Fig. 7(a), we observe that the correlation dimension of this network is approximated to the Hausdorff dimension of the Sierpinski gasket, i.e., $ln3/ln2$ [27]. Correlation sums of series of the multidimensional vectors from the Sierpinski gasket corresponding to the maximal h eigenvalues are presented in Fig. 7(b). The curves for correlation dimension converge to 1.57 when h is more than 3. Again, the fractal network and its transformed time series have the same fractal dimension.

Different from the previous networks with a fixed number of links for each node, a random network is constructed based on a prescribed probability ρ determining whether one node is connected with others. For a random network of size n, the probability of two nodes with the shortest path length less than l approximates to $1 - e^{-1/n(n\rho)^l}$ [28]. For the given l, the average number of nodes centered at a seed node approximates to $C(l) = (n-1)(1 - e^{-1/n(n\rho)^l})$. The first-order Taylor expansion of the previous equation is $(n-1)/n(n\rho)^l$ according to $e^{-1/n(n\rho)^l} \approx 1 - \frac{1}{n}(n\rho)^l$. This shows that the logarithm of an average number of nodes exhibits a linear function with respective to l, i.e., $\log C(l) \approx l \log(n\rho)$, which suggests that a random network has an

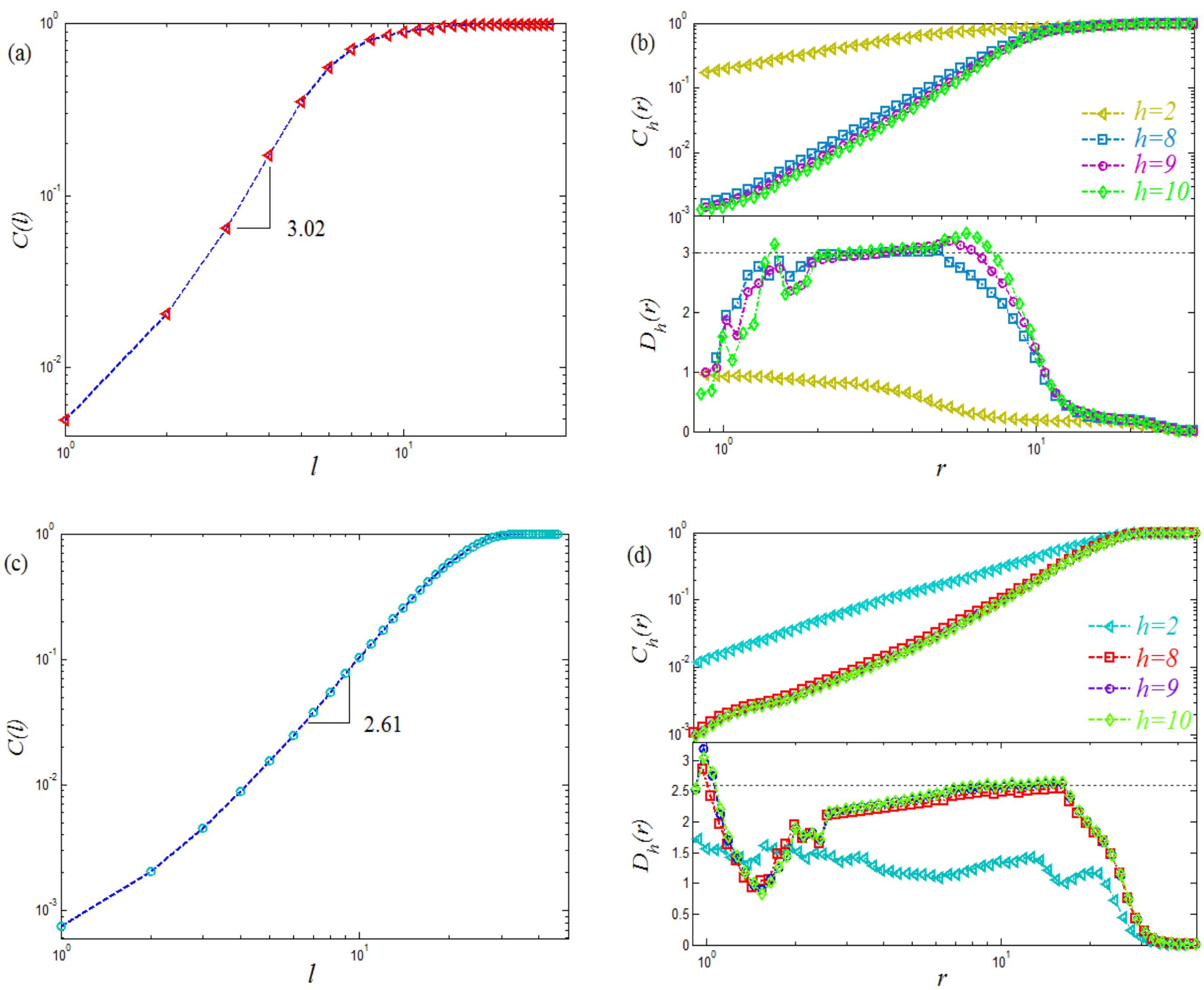

FIG. 9. (Color online) (a) Scaling for the protein-protein interaction network and the correlation dimension is estimated to 3.02. (b) (Top panel) Log-log plot of correlation sum $C_h(r)$ as a function of a distance scalar r for the protein-protein network. (Bottom panel) Local slopes of the correlation sums shown in the upper panel. The correlation dimension is an estimate of $D = 3 \pm 0.05$. (c) Scaling for the electrical power grid network and the correlation dimension is estimated to 2.61. (d) (Top panel) Log-log plot of correlation sum $C_h(r)$ as a function of a distance scalar r for the power grid network. (Bottom panel) Local slopes of the correlation sums shown in the upper panel. The correlation dimension is an estimate of $D = 2.58 \pm 0.05$. The convergent value is marked by a horizontal dashed line.

infinite correlation dimension. We notice that in Fig. 8 correlation sums of series of the multidimensional vectors from a random network fail to converge to a fixed scaling. Therefore, the multidimensional vectors transformed from a random network also give an infinite correlation dimension. This result implies that the dynamics described by the multidimensional vectors of a random network may correspond to a stochastic system.

The structural relationship of the previous synthetic networks is mapped to the spatial relationship coordinated by the multidimensional vectors from these networks obtained by CMDS. Furthermore, geometrical features related to each network are still approximately preserved during transformations. Therefore, we get geometrical invariability in terms of measure of correlation dimension. As a further validation, we examine two real networks. One is the protein-protein interaction network having 1609 nodes and 5546 links [29], which has been confirmed as a fractal network [14]. The other is the electrical power grid of the western United States, with 4941 nodes and 6594 edges [26]. From Fig. 9 we observe an intermediate scaling regime and a scaling $C(l) \sim l^{\alpha}$ ($\alpha = 3.02$) for the first network. This exponent value is a little lower than the box-counting dimension of this network [14], since the box-counting dimension is an upper boundary of Hausdorff dimension [16]. The scaling for the multidimensional vectors shows $C_h(r) \sim l^{\beta}$ ($\beta = 3$). This network and its multidimensional space representation almost have the same correlation dimension. The estimation of correlation dimension of the latter network is consistent with the dimension of this network suggested in Ref. [24]. We also notice the consistency of correlation dimensions for another network and its multidimensional vectors in the corresponding scaling regimes.

V. IDENTIFICATION OF DYNAMICAL TRANSITION THROUGH TRANSFORMATION

Here we demonstrate the dual characterization of complex systems by means of a quasi-isometric transformation. We use the logistic map given by $x_{n+1} = ax_n(1 - x_n)$, where a is the control parameter. Figure 10(a) shows the bifurcation diagram versus $a \in [3.5,4]$ with a step size of 0.007 5, where the logistic map exhibits various dynamical behaviors [30]. Here, we characterize it in terms of network measurement. For each value of a, we generate 1000 data from the logistic map and choose the threshold ε that satisfies the condition of the quasi-isometric transformation. We calculate the clustering coefficient C from the transformed networks. A transition of dynamical behaviors is exactly captured by this network statistics. In particular, the periodic behavior corresponds to the maximal value of C (i.e., $C = 1$), whereas smaller values of C (i.e., less than 0.82) are calculated from the clearly chaotic regime. Some other behaviors such as accumulation point and interior crisis may result in intermediate values of C [30]. Results imply that a dynamically equivalent transformation ensures that dynamical behaviors hidden in a time series can be strictly preserved and then captured by network statistics. Significantly, the quasi-isometric transformation theorem makes it possible to investigate complex systems from both perspectives and thereby achieves a comprehensive understanding for them. Based on such a transformation, more network-based methods may be proposed to describe specific dynamical behaviors of complex systems as well as their characteristic parameters that have been studied by time-series methods such as Lyapunov exponent or recurrence times [31,32].

VI. DISCUSSION ON FRACTIONAL BROWNIAN PROCESS

We have so far illustrated our findings with deterministic systems. Of course, the quasi-isometric theorem is also applicable to stochastic processes. To address this case, we choose the fractional Brownian motion (fBm) with Hurst exponents H ranging from 0.5 to 0.95. We repeat the previous experimental procedure with these fractional Brownian motions and find that the correlation dimension D of their networks stays almost the same as the original fractal dimension of a stochastic

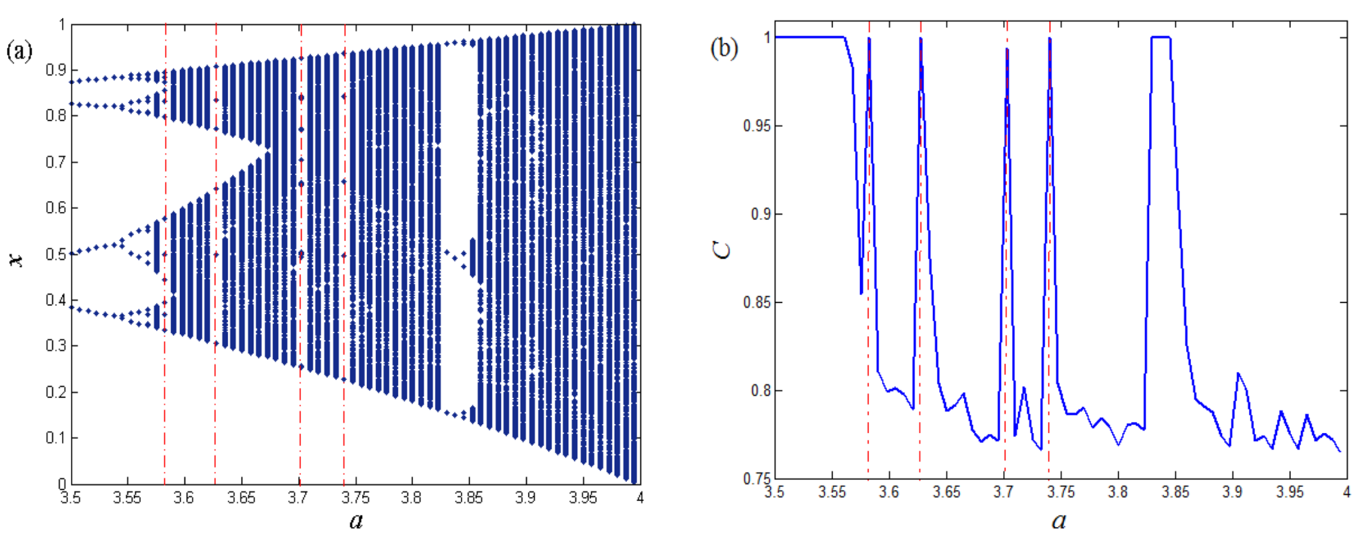

FIG. 10. (Color online) (a) Bifurcation diagram of the logistic map versus the parameter a. (b) Clustering coefficient C of the transformed networks.

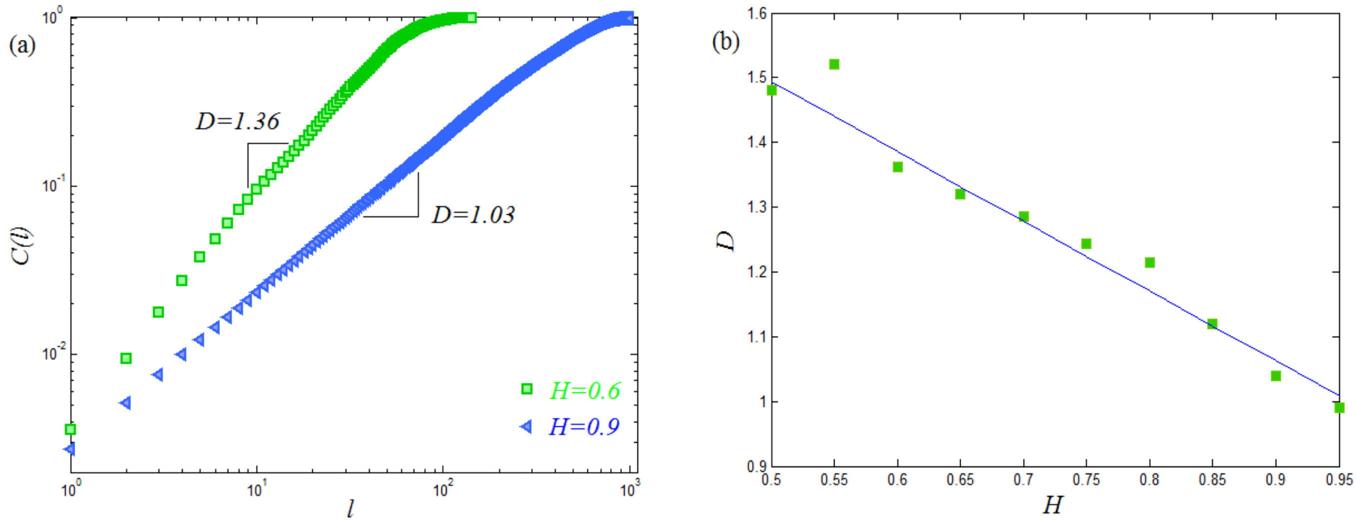

FIG. 11. (Color online) (a) Correlation sum $C(r)$ for the transformed network from the fBm versus the shortest path length l. (b) Relationship between the Hurst exponent H of a series of original fBm data and a correlation dimension D of their transformed networks.

time series. Typical results are plotted in Fig. 11(a). More significantly, the dimension estimated from these networks contains an approximately linear function with respect to the Hurst exponent H ($D = -1.07 * H + 2.03$), which indicates the existing relationship between the Hurst exponent H and fractal dimension d for the fBm ($d = 2 - H$) [33]. We notice that when the Hurst exponent decreases to near 0.5, the stochastic data is less correlated and there is more stochastic fluctuation. It gradually becomes difficult to exactly preserve the original distance during transformation. For the fBm with a Hurst exponent lower than 0.5, due to anticorrelation it is almost impossible to achieve a proper threshold value in practical simulation so as to ensure geometrical perseveration over transformation.

VII. CONCLUSION

In this work, we give theoretical evidences for the geometrical invariability of transformations so that a time series and its network representation (or vice versa) can be dynamically equivalent. This conclusion is based on a theorem of quasi-isometry, which implies that underlying geometrical features of complex systems are preserved. As a result, the fractal dimension stays the same in a justified quasi-isometric transformation. Our theorem is applicable to the transformation methods proposed in Refs. [7,34].

The correlation dimension of a network has been described in Ref. [16]. In that article, a trajectory of random walks on a network as a time series is reconstructed with embedding dimensions. The correlation dimension estimated on the reconstructed phase space is regarded as a network correlation dimension. In contrast to this method, we directly transform a network into multidimensional vectors by the technique of classical multidimensional scaling. As the output vectors are an optimal configuration of the input network structure, CMDS enables a quasi-isometric transformation. The distancelike information of a network is preserved during such a transformation. Consequently, the correlation dimension of the multidimensional vectors exactly reflects the network dimension. Numerical analysis of various synthetic and real networks confirms this finding. One advantage of our method is that the quasi-isometric theorem provides a theoretical base for identifying network dimensions. The numerical results of more networks support our conclusion.

The quasi-isometric transformation theorem (i.e., geometrical invariability) serves as a fundamental proof that ensures an equivalent characterization of complex systems from the dual perspective of a time series and a network. We have so far illustrated the feasibility and utility of a network-based approach for identifying the underlying bifurcation of a time series given such transformation. Our theorem suggests that by means of an equivalent transformation it is feasible to explore the dynamical characteristics of the complex system from a new perspective or even mine extra information that is uncovered from the present perspective. Consideration is then shifted to how to extract effective network (or time-series)-based methods for this purpose.

ACKNOWLEDGMENTS

We acknowledge support from the National Science Foundation Committee (NSFC) of China under Project No. 60801014, and a Fundamental Research Project of Shenzhen under Projects No. JC201005260116A and No. JCYJ20120613144110654.

[1] R. V. Donner, M. Small, J. F. Donges, N. Marwan, Y. Zou, R. X. Xiang, and J. Kurths, Int. J. Bifur. Chaos **21**, 1019 (2010).
[2] J. Zhang and M. Small, Phys. Rev. Lett. **96**, 238701 (2006).
[3] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuño, Proc. Natl. Acad. Sci. USA **105**, 4972 (2008).
[4] M. Thiel, M. C. Romano, and J. Kurths, Phys. Lett. A **330**, 343 (2004).

[5] A. S. L. O. Campanharo, M. I. Sirer, R. D. Malmgren, F. M. Ramos, and L. A. N. Amaral, PLoS ONE **6**, e23378 (2011).
[6] Z. K. Gao, N. D. Jin, W. X. Wang, and Y. C. Lai, Phys. Rev. E **82**, 016210 (2010).
[7] R. V. Donner, Y. Zou, J. F. Donges, N. Marwan, and J. Kurth, New J. Phys. **12**, 033025 (2010).
[8] G. McGuire, N. B. Azar, and M. Shelhamer, Phys. Lett. A **237**, 43 (1997).
[9] Y. Hirata, S. Horai, and K. Aihara, Eur. Phys. J. Special Topics **164**, 13 (2008).
[10] X. F. Liu, C. K. Tse, and M. Small, Physica A **389**, 126 (2010).
[11] W. N. Goncalves, A. S. Martinez, and O. M. Bruno, Chaos **22**, 033139 (2012).
[12] Y. Shimada, T. Ikeguchi, and T. Shigehara, Phys. Rev. Lett. **109**, 158701 (2012).
[13] G. Robinson and M. Thiel, Chaos **19**, 023104 (2009).
[14] C. M. Song, S. Havlin, and H. A. Makse, Nature (London) **433**, 392 (2005).
[15] V. M. Eguiluz, E. Hernandez-Garcia, O. Piro, and K. Klemm, Phys. Rev. E **68**, 055102 (2003).
[16] L. Lacasa and J. Gómez-Gardeñes, Phys. Rev. Lett. **110**, 168703 (2013).
[17] P. Grassberger and I. Procaccia, Phys. Rev. Lett. **50**, 346 (1983).
[18] M. Cencini, M. Falcioni, E. Olbrich, H. Kantz, and A. Vulpiani, Phys. Rev. E **62**, 427 (2000).
[19] J. B. Gao, J. Hu, W. W. Tung, and Y. H. Cao, Phys. Rev. E **74**, 066204 (2006).
[20] J. Wang, *Geometric Structure of High-Dimensional Data and Dimensionality Reduction* (Springer Press, New York, 2011).
[21] I. J. Schoenberg, Transactions of the American Mathematical Society **44**, 522 (1938).
[22] P. W. Nowak and G. Yu, *Large Scale Geometry, EMS Textbooks in Mathematics* (European Mathematical Society Publishing House, Zurich, Switzerland, 2012).
[23] J. Feder, *Fractals* (Plenum Press, New York, 1988).
[24] L. Guo and X. Cai, Chin. Phys. Lett. **26**, 088901 (2009).
[25] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge, UK, 2004).
[26] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).
[27] J. S. Kim, K.-I. Goh, B. Kahng, and D. Kim, Chaos **17**, 026116 (2007).
[28] A. Fronczak, P. Fronczak, and J. A. Holyst, Phys. Rev. E **70**, 056110 (2004).
[29] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási, Nature (London) **407**, 651 (2000).
[30] R. Wackerbauer, A. Witt, H. Atmanspacher, J. Kurth, and H. Scheingraber, Chaos, Solitons Fractals **4**, 133 (1994).
[31] J. B. Gao, Phys. Rev. Lett. **83**, 3178 (1999).
[32] J. B. Gao and H. Q. Cai, Phys. Lett. A **270**, 75 (2000).
[33] J. B. Gao, Y. H. Cao, W. W. Tung, and J. Hu, *Multiscale Analysis of Complex Time Series: Integration of Chaos and Random Fractal Theory, and Beyond* (Wiley-Interscience Press, New York, 2007).
[34] X. K. Xu, J. Zhang, and M. Small, Proc. Natl. Acad. Sci. USA **105**, 19601 (2008).