PHYSICAL REVIEW E **86**, 026203 (2012)

Estimating network topology by the mean first-passage time

Pu Yang,[1] Qun Wang,[1,2] and Zhigang Zheng[1,*]

[1]*Department of Physics and the Beijing–Hong Kong–Singapore Joint Center for Nonlinear and Complex Systems (Beijing), Beijing Normal University, Beijing 100875, People's Republic of China*

[2]*School of Physics and Electronic Engineering, Xuzhou Normal University, Xuzhou, Jiangsu 221116, People's Republic of China*

(Received 21 November 2011; revised manuscript received 22 June 2012; published 13 August 2012)

In this work, we employed the concept of the first-passage time in stochastic processes to estimate node degrees and the degree distribution of a network. A statistical exploration of the coupling reveals the relation between the node degree and the coupling term. In practical terms, an effective way to reveal the statistical property is to investigate the differences between coupled oscillators in a network and uncoupled ones with the same initial states. We discovered a monotonically decreasing relation between the node degree and the mean first-passage time (MFPT) for the evolution of the coupled node deviating from the uncoupled one. Moreover, this relation can be understood as the competition of different relaxational time scales. The MFPT method is independent of both the dynamics of the nodes and the topological properties of the network. This might be advantageous in our efforts to build a bridge between the topological property and the dynamics of a network.

DOI: 10.1103/PhysRevE.86.026203 PACS number(s): 89.75.−k, 05.45.Tp

I. INTRODUCTION

The studies of complex networks began with the explorations of small-world [1] and scale-free [2] networks at the end of the 1990s, and they have continued for more than a decade since then. Many systems can be considered as networks, including the World Wide Web, transportation systems, telephone networks, and so on. These types of networks can be designed manually and manipulated. The dominant studies on these networks have been focused on the statistical properties of the given networks or the effects of topologies on the network dynamics, especially synchronization. On the other hand, many natural networks are generated in a self-organized manner, including gene networks, climate networks, neural networks, ecological networks, etc. [3–5]. The exploration of the topological properties of these networks is a significant aspect of understanding the complexity and organization processes of networks. Revealing the topology of networks may help us to learn the relationship between these topologies and their functioning.

As an important inverse problem, estimating the topologies of networks through time series has become a hot topic in recent years. Many methods have been proposed for estimating network topologies [6–11]. Most of these methods have focused on identifying the existence and distributions of links in a network, or evaluating the strength of some of these links. These methods usually require a very long computational time, which is proportional to N^2, where N is the number of nodes in a network. When the size of a network N is very large, all of these previously proposed methods are inappropriate in practice. In fact, for a very large network, it is not necessary to estimate the complete information of the topology in many circumstances. Several studies have been performed based on this consideration. Bu *et al.* [12] employed the return map to estimate the degree distributions in networks. The return map approach requires no explicit knowledge of dynamics, which is adaptive to many chaotic systems; the only requirement is that the network should be sparse. However, theoretical support is still lacking for this approach. Shen *et al.* [13] also developed a method to estimate the degree distributions of a network of coupled neurons. The amplitude of the action potential was employed. The amplitude method is explained in a local mean-field way. However, this method is only applicable to some neuron models in burst synchronization states. Liang *et al.* [14] proposed a method that can predict many statistical characteristics of a network with specific dynamics just by adding a weak signal on one node in the network. The basic idea behind these methods is to find an appropriate quantity that has a monotonic relation with one aspect (i.e., the degree) of the network, and to use this parameter to uncover the value or the distribution of the character in the network.

In fact, looking for an appropriate dynamical index that is universally connected to topological properties of networks is still a challenging and open issue. This challenge is at the heart of the dynamical identification of network topology. In this paper, we propose a method for finding a parameter to estimate the relative values of degrees in a network. This approach borrows an important concept in stochastic process [15], i.e., the mean first-passage time (MFPT), which is defined here as the time that the dynamical state of the network evolution deviates from the isolated auxiliary system to a threshold. It is found that a good relation can be established between the node degree and the MFPT. Our statistical analysis establishes a universal correspondence between the coupling term and the degree, which is independent of the specific form of node dynamics and the form of the output function. The MFPT gives the time scale describing the deviation of the coupled state from the isolated state. Further theoretical investigations indicate that the MFPT is closely related to the maximum Lyapunov exponent that labels the node degree. Therefore, the theoretical relation we obtain implies a competition between different time scales. Our theoretical results agree well with numerical simulations.

*zgzheng@bnu.edu.cn

II. THEORETICAL ANALYSIS

A. Statistical analysis: The coupling term

The dynamics of a network composed of N identical oscillators can be written as follows:

$$\dot{\mathbf{x}}_i = \mathbf{F}(\mathbf{x}_i) - s\sum_{j=1}^{N} L_{ij}\mathbf{H}(\mathbf{x}_j), \quad i = 1,2,\ldots,N, \tag{1}$$

where $\mathbf{x}_i = (x_{i1},x_{i2},\ldots,x_{im})^T \in \mathbb{R}^m$ is the m-dimensional state vector of the ith oscillator. The function $\mathbf{F}$, which governs the dynamics of isolated oscillators, is a vector field $\mathbb{R}^m \to \mathbb{R}^m$, and $\mathbf{H}$ is an output function of every oscillator. s is the global coupling strength. The Laplacian matrix $\mathbf{L} = (L_{ij})_{N\times N}$ describes the network topology. $L_{ij} = -1$ if node j is connected to node i; otherwise, $L_{ij} = 0$, and the diagonal element $L_{ii} = -\sum_{j=1,j\neq i}^{N} L_{ij}$. The value of L_{ii}, which corresponds to the input degree k_i of the ith node, is exactly what is to be estimated in this paper.

The dynamics described by Eq. (1) has been investigated extensively in recent years [8,16–18]. In this system, the interactions among nodes are identical, and the couplings among nodes are negative feedback. Due to these properties, if there exists global synchronization among nodes, i.e., $\{\mathbf{x}_j = \mathbf{x}_1 \mid j = 2,3,\ldots,N\}$, the coupling term $\sum_{j=1}^{N} L_{ij}\mathbf{H}(\mathbf{x}_j) = 0$, implying that the information about the topology of the network is lost. Therefore, synchronization is an obstacle in detecting network topology [19]. The value of the coupling strength s is chosen from the weak region where synchronization is unstable.

It is important to investigate the statistical properties of the coupling terms in Eq. (1), since this may give us clues that will help us to understand the topological features of the network. Before the statistical analysis of the coupling term, let us first introduce some results in probability theory. Assume that two independent stochastic time series $x(t)$ and $y(t)$ possess the same statistical distribution $P(E,\sigma)$, where E and σ are the expectation and the standard deviation of the distribution, respectively. Then the linear statistical distribution of the combination of $x(t)$ and $y(t)$ obeys

$$ax(t) + by(t) \sim P[(a+b)E,\sqrt{a^2+b^2}\sigma], \tag{2}$$

where $a,b \in \Re$. Specifically, when $b = 0$, we obtain

$$ax(t) \sim P(aE,|a|\sigma). \tag{3}$$

The coupling term provides a significant measure integrating the information of connectivity. Our work starts with an analysis of the coupling term by separating the diagonal term from the coupling part of the equations of motion, i.e.,

$$\sum_{j=1}^{N} L_{ij}\mathbf{H}(\mathbf{x}_j) = L_{ii}\mathbf{H}(\mathbf{x}_i) + \sum_{j=1,j\neq i}^{N} L_{ij}\mathbf{H}(\mathbf{x}_j). \tag{4}$$

The ith diagonal element of the adjacent matrix is associated with the degree (in-degree) of the ith node, i.e., $L_{ii} = k_i$. Equation (4) indicates that one could consider the linking information in the coupling term by separately estimating the contributions of the diagonal and off-diagonal parts given by the two terms on the right-hand side in Eq. (4). Assume that the output time series have the same probability distribution,

$$\mathbf{H}(\mathbf{x}_i) \sim P_c(\mathbf{E},\boldsymbol{\sigma}), \quad i = 1,2,\ldots,N. \tag{5}$$

The m-dimensional vectors $\mathbf{E}$ and $\boldsymbol{\sigma}$ are the long-time or ensemble average and the standard deviation of the distribution, respectively. For the first term, according to the relation (3) and noticing that $L_{ii} = k_i$, we obtain

$$L_{ii}\mathbf{H}(\mathbf{x}_i) \sim P_c(k_i\mathbf{E},k_i\boldsymbol{\sigma}). \tag{6}$$

For the off-diagonal term on the right-hand side of Eq. (4), we wish to apply the relation (2) to discuss its statistical property. Suppose that the terms $\mathbf{H}(\mathbf{x}_j)$ for different j in the summation of the second term are independent of each other. According to the relation (2), we obtain the distribution of the second term of Eq. (4) as

$$\sum_{j=1,j\neq i}^{N} L_{ij}\mathbf{H}(\mathbf{x}_j) \sim P_c(-k_i\mathbf{E},\sqrt{k_i}\boldsymbol{\sigma}). \tag{7}$$

According to relations (2), (6), and (7), we obtain the distribution of the coupling term,

$$\sum_{j=1}^{N} L_{ij}\mathbf{H}(\mathbf{x}_j) \sim P_c\big(0,\sqrt{k_i^2+k_i}\boldsymbol{\sigma}\big). \tag{8}$$

The above results are obtained on the basis of the statistical independence of output time series $\mathbf{H}(\mathbf{x}_j)$. However, $\mathbf{H}(\mathbf{x}_j)$ for different j are not strictly independent of each other as long as the coupling strength $s \neq 0$. Let us take the one-dimensional situation as an example. Denote the correlation $\sigma_{ij} = \langle [H(x_i) - \langle H(x_i)\rangle][H(x_j) - \langle H(x_j)\rangle]\rangle$. According to the relation (5) in one dimension, we have

$$\sum_{j=1,j\neq i}^{N} L_{ij}H(x_j) \sim P_c\left(-k_iE,\sigma\sqrt{k_i + \hat{\sum} L_{ij}L_{il}\frac{\sigma_{jl}}{\sigma^2}}\right), \tag{9}$$

where $\hat{\sum}$ means $\sum_{j,l=1;j,l\neq i;j\neq l}^{N}$.[1] There are $k_i(k_i-1)$ nonzero terms in the $\hat{\sum}$. We find that the correlation coefficients $\{\sigma_{jl}/\sigma^2 | j,l = 1,2,\ldots,N\}$ obey Gaussian distribution, with zero expectation and small standard deviation (<0.1) in the weak-coupling condition. This has been proven in numerical simulations for different cases, such as different kinds of oscillators, output functions, and topologies of networks. This symmetry property of the distribution guarantees that the term $\hat{\sum}$ is small for large degree, and the small standard deviation guarantees that the term $\hat{\sum}$ is small for small degree. Therefore, the contribution of the correlations to the deviation in (9) can be neglected, and the relations (7) and (8) still hold approximately in the weak-coupling condition.

[1][In the one-dimensional condition, the contribution of the correlation between output time series to the coupling terms can be worked out: $\langle[\sum_{j=1,j\neq i}^{N} L_{ij}H(x_j) - \langle\sum_{j=1,j\neq i}^{N} L_{ij}H(x_j)\rangle]^2\rangle = \sum_{j=1,j\neq i}^{N}\langle[L_{ij}H(x_j)]^2\rangle - \sum_{j=1,j\neq i}^{N} L_{ij}^2\langle H(x_j)\rangle^2 + \hat{\sum}L_{ij}L_{il}\langle H(x_j)H(x_l)\rangle - \hat{\sum}L_{ij}L_{il}\langle H(x_j)\rangle\langle H(x_l)\rangle = \sum_{j=1,j\neq i}^{N} L_{ij}^2\sigma^2 + \hat{\sum}L_{ij}L_{il}\sigma_{jl} = \sigma^2(k_i + \hat{\sum}L_{ij}L_{il}\frac{\sigma_{jl}}{\sigma^2})$.]

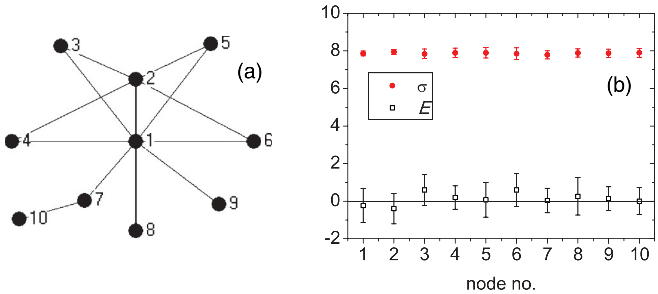

FIG. 1. (Color online) (a) The topology of a 10-node network. The largest degree of nodes is 8 ($i = 1$). Five nodes labeled by $i = 3$, 4, 5, 6, and 7 have the same degree $k_i = 2$, and four of these five, $i = 3$, 4, 5 and 6, are topologically symmetric. (b) E (open square) and σ (filled circle) are the first components of the expectation and the standard deviation of normalized coupling terms $(k_i^2 + k_i)^{-1/2} \sum_{j=1}^{N} L_{ij} \mathbf{H}(\mathbf{x}_j)$. The error bars come from the samples with different initial conditions. The deviations σ are almost the same, independent of the degrees of nodes. The expectations E are close to 0 (the solid line). All these results agree well with our analytical prediction (8). $s = 0.01$.

The relation (8) indicates that the statistical distribution of the coupling fluctuation is a zero-mean distribution with a variance related to the node degree. For different nodes, the fluctuation due to the coupling obeys the same statistical distribution. The nodes with the same degree possess the same distribution, and nodes with different degrees have different widths of the distribution. Therefore, we have established a correspondence between the coupling term and the degrees of the nodes. The standard deviation for different nodes corresponds to the diffusion coefficient. This relation is universal. It is independent of the specific form of the node dynamics $\mathbf{F}$ and the form of the output function $\mathbf{H}$.

Now let us perform numerical simulations to test the validity of the relation (8). Consider a network with the topology given in Fig. 1(a) and node dynamics described by chaotic Lorenz oscillators:

$$\begin{aligned} F_1(\mathbf{x}) &= 10(x_2 - x_1), \\ F_2(\mathbf{x}) &= 28x_1 - x_1x_3 - x_2, \\ F_3(\mathbf{x}) &= x_1x_2 - \tfrac{8}{3}x_3. \end{aligned} \tag{10}$$

The output function $\mathbf{H}(\mathbf{x})$ is assumed to be linear:

$$\mathbf{H}(\mathbf{x}) = (x_1; 0; 0). \tag{11}$$

In numerical simulations, the fourth-order Runge-Kutta algorithm is employed with the time step size $\delta t = 0.01$. In Fig. 1(b), it can be found that $E \approx 0$, and the standard deviation σ's are the same for all nodes, verifying the validity of our analytical prediction (8). In this system, when $s > 0.16$, partial synchronization becomes stable according to the master-stability function [20]. And when $s < 0.1$, the standard deviation of correlation coefficients σ_{ij}/σ^2 is less than 0.1.

B. Statistical analysis: The first-passage time approach

A pertinent problem in building the bridge between the coupling effect and the degree of nodes is how to extract this information in practice. The above discussions focus on the analysis of coupling terms. However, it is much easier to get the time series $\mathbf{x}_i(t)$ than the coupling term $\sum_{j=1}^{N} c_{ij} \mathbf{H}(\mathbf{x}_j)$ in practice. A natural idea in utilizing the information of the time series is to set isolated reference nodes, whose dynamics are the same as those in the network, and make a comparison of the dynamics between the coupled and isolated cases. Let us set N auxiliary isolated oscillators $\mathbf{y}_i(t), i = 1,2,\ldots,N$, governed by

$$\dot{\mathbf{y}}_i = \mathbf{F}(\mathbf{y}_i), \tag{12}$$

which are identical to the ones in the network when the global coupling strength $s = 0$. Let these uncoupled oscillators evolve from the same initial states as those in the network. Obviously the evolution of the isolated node $\mathbf{y}_i(t)$ will deviate from the orbit given by the coupled node $\mathbf{x}_i(t)$. One can thus introduce the difference between these two states,

$$e_i(t) = \|\mathbf{y}_i(t) - \mathbf{x}_i(t)\|, \tag{13}$$

to measure the deviation of the time series $\mathbf{y}_i(t)$ from $\mathbf{x}_i(t)$. Due to the chaotic motion and asynchronous property of the system, the evolution of the coupling term exhibits a stochastic-like behavior. Moreover, the bigger the standard deviation of the coupling term (corresponding to the diffusion rate), the faster $e(t)$ grows. Statistically, an important quantity, i.e., the first-passage time (FPT), can be introduced to measure the growth rate of $e(t)$. It is interesting to reveal the relationship between the FPT and the degree indirectly. According to the definition of the FPT in [15], we can define it here as the time at which the time series $e(t)$ passes a given threshold θ for the first time. Obviously the FPT is also a stochastic quantity, and then the MFPT T_1 can be computed by averaging the FPTs obtained from an ensemble of different initial states $\{\mathbf{x}(t_0^i) \mid i = 1,2,\ldots\}$. Assume the MFPT is related to the degree by satisfying the following function:

$$T_1 = G(s'k), \tag{14}$$

where s' is a scaling coefficient. Is the function G monotonic? What is the form of the function G? These questions are important issues in this paper. If G is monotonic, the relative value of the degrees can be established. Furthermore, if the form of G can be determined, we can make a transformation $T_1' = G^{-1}(T_1) = s'k \propto k$. The distributions of T_1' and k are the same [12,13], therefore the parameter s' is irrelevant.

In the following, we first choose a small network with 10 nodes [the topology is shown in Fig. 1(a)] and concentrate on building a correspondence between the MFPT (dynamical properties) and the degree (topological feature). The results obtained in a small network can be naturally generalized to large networks with large numbers of nodes (e.g., $N = 1000$).

III. NUMERICAL RESULTS

A. Topology identification of a small network

The topology of the network given by Fig. 1(a) is interesting because there are various kinds of nodes in the network, such as nodes with different degrees, and topologically symmetric and asymmetric nodes with the same degree. Therefore, the dependence of T_1 on the degree of nodes that are topologically symmetric can be well examined. Denote the probability distribution of FPT by P_t. In Fig. 2(a), the FPT distribution P_t

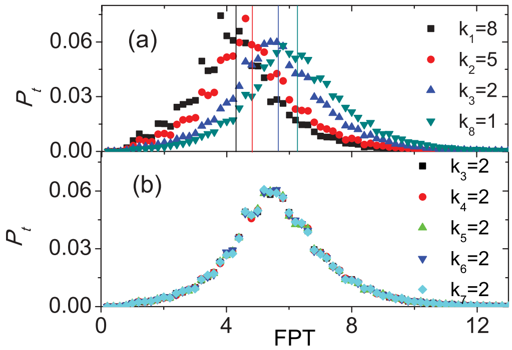

FIG. 2. (Color online) The probability distribution of the first-passage time for different nodes in the 10-node network. (a) The distribution P_t for four nodes with different $k_1 = 8, k_2 = 5, k_3 = 2,$ and $k_8 = 1$. The solid lines label the average values of FPT, T_1. (b) The distribution P_t for five nodes with the same degree $k = 2$. It can be observed that these five curves coincide with each other.

for different nodes is plotted. It can be found that the MFPT T_1 (labeled as the position of the average value) for nodes with different degrees is obviously different. Moreover, for nodes with larger degrees, the position of the distribution peak shifts to the left, implying that the MFPT decreases monotonically upon increasing the degree.

In Fig. 2(b), the FPT distribution P_t for nodes $i = 3, 4, 5,$ 6, and 7 is given. These five nodes possess the same degree $k = 2$ but different topological symmetries. It can be clearly found that all P_t's coincide with each other, indicating that the FPT distribution for nodes with the same degree is the same. Therefore, the MFPT T_1 only depends on the degrees of the nodes, and it is independent of the topological symmetry in a network.

Another critical problem is the effect of the threshold θ on the MFPT T_1. Surely MFPT T_1 will be shorter for a smaller threshold θ. It is important in practical analysis to choose an appropriate value of the threshold. In Fig. 3, we plot the dependence of the MFPT T_1 on the values of the threshold θ for nodes with different degrees. It can be observed that the curve $T_1 \sim \theta$ is smooth. It can also be found that T_1 is

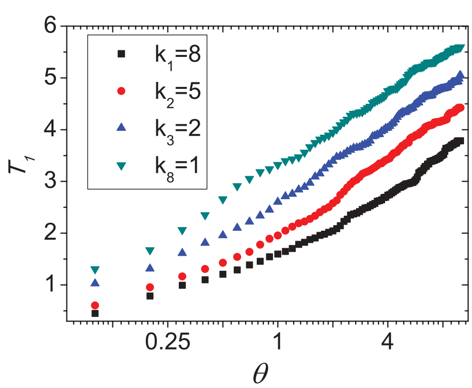

FIG. 3. (Color online) The MFPT against the threshold θ. An approximately logarithmic relation can be found, i.e., $T_1 \propto \ln(\theta)$. $k = 1$ (square), 2 (ball), 5 (up-triangle), and 8 (down-triangle). The relation between T_1 and k is not affected for different values of θ in a large range.

approximately proportional to $\ln\theta$, and the T_1 intervals among curves at different θ's are almost the same. This implies that one can choose the value of the threshold in a wide range. For the network we are discussing here, we can choose the value of $\theta \in (0.5,10)$. In our discussions below, we take $\theta = 1.5$ for Lorenz node dynamics.

B. Topology identification of large networks

It is significant to apply the above MFPT method to large networks. Let us consider a BA network [2] composed of 1000 nodes with degrees ranging from $k_{\min} = 3$ to $k_{\max} = 115$, and the average degree $\langle k\rangle = 6$. We use the label BA1000 in the following discussions. We still adopt the node dynamics as the Lorenz oscillators governed by Eqs. (10) and (11). The dependence of the MFPT T_1 on the degree k is shown in Fig. 4(a) for two different coupling strengths, $s = 0.001$ and 0.002. It can be found that a linear relation is established between T_1 and $\ln k$. Moreover, these two lines are parallel to each other. Therefore, one can replace $\ln k$ by $\ln sk$ to normalize two lines. As shown in Fig. 4(b), the two lines are overlapped with the same slope. This supports our analysis in Eq. (8). Figure 4 reveals that $T_1 = a\ln(sk) + b = a\ln(s'k) = G(s'k)$. Let $T_1' = G^{-1}(T_1) = e^{T_1/a} = s'k \propto k$. In Fig. 4(c), we plot the distribution of T_1'. To make a comparison, the degree distribution $P(k)$ is also given. It can be found that the distribution of T_1' coincides well with the degree distribution for the BA1000 network. Therefore, the MFPT approach is efficient in estimating the degree distribution.

It is very significant to confirm the universality of the form of the function G. We have tested various forms (both linear and nonlinear) of the output function $\mathbf{H}$. It can be found that for different coupling schemes, the G functions are the same, $G(s'k) = a\ln(s'k)$, where $a \approx -1.02 \pm 0.09$. Just two examples are shown in Fig. 5(a).

It is also interesting to study the influence of node dynamics on the G function. We have tested the MFPT method for other kinds of chaotic oscillators. For the case of Chua's circuit oscillators, the result is similar to the Lorenz node dynamics with a logarithmic relation and $a \approx -2.25$. But for the case of Rössler oscillators, Fig. 5(b) shows that $\ln T_1 = \gamma\ln(sk) + b = \gamma\ln(s'k)$, $T_1 = G(s'k) = (s'k)^\gamma$, which is much different from the case of Lorenz dynamics. An explanation for this difference will be given in the following discussions.

Moreover, we have tested the effects of different kinds of networks (e.g., WS networks [1] and directional ER networks [21]). Our extensive studies reveal that the form of G is not affected by the topology of the networks.

C. MFPT-based network identification for unknown dynamics

The above discussions of the MFPT approach requires explicit knowledge of the node dynamics and the coupling function. However, in many cases one can only get the time series $\{x_i(t)\}$. Therefore, it is important to study whether the present MFPT approach can be applicable in cases in which there is no knowledge of the node dynamics and the coupling functions. Here we show that the MFPT approach can be naturally generalized to this case.

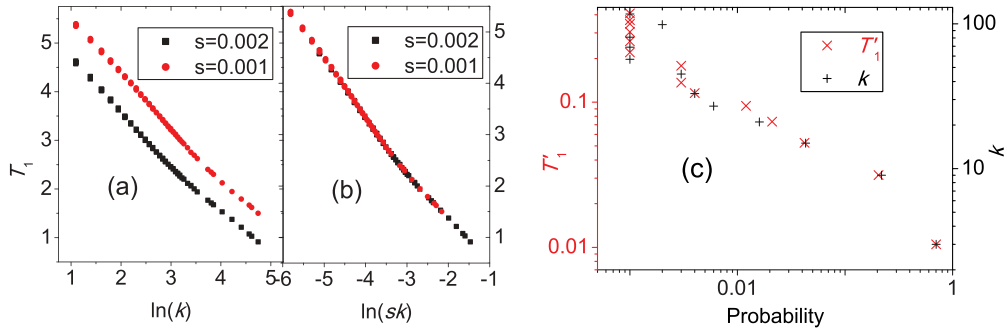

FIG. 4. (Color online) (a) The relationship between the MFPT T_1 and the degree k for a BA network with 1000 Lorenz nodes for different coupling strengths $s = 0.001$ and 0.002. Two curves are parallel. (b) Two parallel lines coincide after a renormalization of ln(k). This implies that T_1 only depends on sk. The slope $a \approx -1.1$. (c) A comparison of the degree distribution (black +) and the distribution of $T_1' = \exp(-T_1)$ (red ×). It can be found clearly that both distributions coincide, indicating the validity of our approach.

To do this, the key point in performing the MFPT approach is to find time series that can replace the auxiliary dynamics $\{\mathbf{y}(t)\}$, which is produced by the isolated oscillator dynamics $\dot{\mathbf{y}} = \mathbf{F}(\mathbf{y})$. A simple idea is to find a node with a small degree (for example, a leaf node with the degree $k = 1$) in the network to generate the auxiliary time series. There may be several methods for discovering such types of nodes. For example, one can perform the following procedure: (i) choose several pairs of nodes; (ii) for each pair, take one node as an auxiliary node, and get the MFPT T_1 of the other node; (iii) take this value as the MFPT T_1 for this node pair. The degrees of the nodes in the pair with larger T_1 are both smaller [as shown in Fig. 6(a)].

Without losing generality, let us assume that the degree of the first node is small. Then the time series $\{\mathbf{x}_1(t)\}$ can be considered as $\{\mathbf{y}_j(t)\}$. Then the time series of the jth node ($j \neq 1$) beginning from $\mathbf{x}_1(t_0^1)$, $\mathbf{x}_j(t_0^j) = \mathbf{x}_1(t_0^1)$ is required. This can be easily achieved by many methods in chaos control [22]. When $\mathbf{x}_j(t_0^j) = \mathbf{x}_1(t_0^1)$, the control should be released. Then one can get the required time series, and the FPT would be fixed. By repeating the above procedure, the MFPT T_1^j can be figured out. It should be emphasized that the system should be in an asynchronous state so that the MFPT approach works, because synchronization is an obstacle in identifying the topology of a network in this approach.

In the above BA1000 network with Lorenz oscillators, we take node 1 (with the degree $k_1 = 3$) as a reference node. Then we compute the MFPT $\{T_1^j \mid j = 2,3,\ldots,1000\}$ and rank them in ascending order [see Fig. 6(b)]. The ideal result should be that the relation between rank and k is monotonically decreasing. Although the node 1 that we choose here is not an absolutely isolated oscillator, the errors shown in Fig. 6(b) are small. If the difference of a pair of degrees $\Delta k > 2$, the rank of the pair of nodes by T_1 is shown to be correct.

IV. DISCUSSIONS

To understand the relationship between the MFPT and the degree, let us analyze the particular details of the node behavior in one sample. The dynamics of $\mathbf{y}_i(t)$ is described by Eq. (12). One can simplify the dynamics of $\mathbf{x}_i(t)$ given by Eq. (1) by separating the coupling term into two parts based on Eq. (7),

$$\sum_{j=1, j\neq i}^{N} L_{ij}\mathbf{H}(\mathbf{x}_j) = -k_i\mathbf{E} + \boldsymbol{\xi}_i,$$

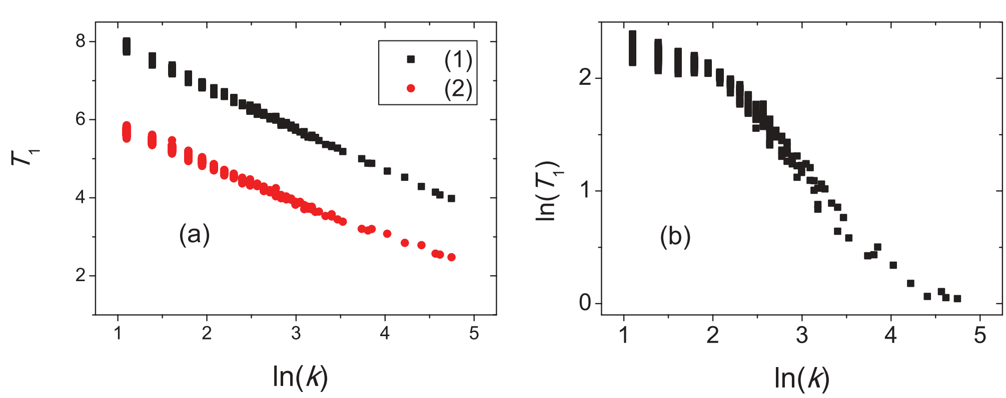

FIG. 5. (Color online) In the BA1000 network. (a) Different output functions are tested with the same Lorenz oscillators. The relationship between the degree and MFPT is also a logarithm. (1) $\mathbf{H}(\mathbf{x}) = (x_1 * x_2; 0; 0)$ (black square) and the slope $a \approx -1.1$; (2) $\mathbf{H}(\mathbf{x}) = (0; 0; x_2)$ (red ball) and the slope $a \approx -0.94$. Curves are logarithmic, with almost the same coefficients. (b) Rössler oscillators $\mathbf{F}(\mathbf{x}) = [-x_2 - x_3; x_1 + 0.2x_2; 0.2 + (x_1 - 10)x_3]$,$\mathbf{H}(\mathbf{x}) = (x_1; 0; 0)$,$\theta = 1$,$s = 0.002$. $G(k) \propto k^\gamma$,$\gamma = -0.97$. This is a power law, in contrast with the case of Lorenz oscillators.

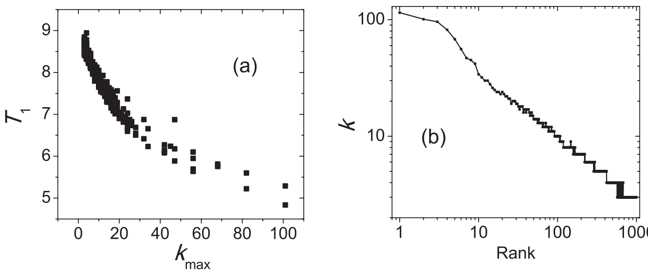

FIG. 6. (a) The maximum degree $k_{\max}$ in each pair of nodes against the MFPT T_1; the big T_1 only maps to a pair of small degree nodes. A total of 1000 pairs of nodes are chosen randomly. (b) The MFPT T_1 of all nodes by using a reference node ($k = 3$) in the same network where the nodes are ranked by T_1. The degree of the first one is the largest.

where $\{\boldsymbol{\xi}_i, i = 1,2,\ldots,N\}$ are independent noise terms with zero mean and the deviation $\sqrt{k_i}\boldsymbol{\sigma}$. $\mathbf{E}$ is independent of the node and the sample because it denotes the long-time or ensemble average of $\mathbf{H}[\mathbf{x}_i(t)]$. Compared with the deviation of the diagonal term $k_i\boldsymbol{\sigma}$ in relation (6), the noise $\boldsymbol{\xi}_i$ is small. Hence we discard the noise term, and Eq. (1) can be simplified as

$$\dot{\mathbf{x}}_i = \mathbf{F}(\mathbf{x}_i) - sk_i[\mathbf{H}(\mathbf{x}_i) - \mathbf{E}]. \tag{15}$$

Equation (15) will be analyzed as the dynamics of $\mathbf{x}_i(t)$ below.

For convenience, we discard the subscript i and redefine the dynamics of the difference between coupled and uncoupled oscillators $e(t)$ as

$$e(t) = \|\mathbf{y}(t) - \mathbf{x}(t)\|_2 = \sqrt{[\mathbf{y}(t) - \mathbf{x}(t)]^2}. \tag{16}$$

Although we take a different form of Eq. (13) as the definition of the difference $e(t)$ in numerical simulations, there is no essential difference in the following analysis given below. All the analytical conclusions are still available in our numerical results. Let us consider the auxiliary trajectory $\mathbf{y}'$, which starts from the initial state $\mathbf{y}'(t-\Delta t) = \mathbf{x}(t-\Delta t)$ and ends at $\mathbf{y}'(t)$ (Fig. 7). The dynamics $\mathbf{y}'(t)$ is governed by Eq. (12). We further introduce two differences $e'(t) = \|\mathbf{y}(t) - \mathbf{y}'(t)\|_2$ and $e''(t) = \|\mathbf{y}'(t) - \mathbf{x}(t)\|_2$. The difference given by Eq. (16) can be approximately considered as the composition of two differences, i.e., $e \approx e' + e''$. According to chaotic theory, $e'(t) \approx e'(t-\Delta t)\exp(\lambda\Delta t)$, where λ is the largest Lyapunov exponent of the isolated node dynamics, which is described by the chaotic dynamics Eq. (12). From the definitions of e' and $\mathbf{y}'$, one gets $e'(t-\Delta t) = \sqrt{[\mathbf{y}(t-\Delta t) - \mathbf{y}'(t-\Delta t)]^2}$. For $\mathbf{y}'(t-\Delta t) = \mathbf{x}(t-\Delta t)$, $e'(t-\Delta t) = e(t-\Delta t)$, which can be seen clearly in Fig. 7. Then one gets

$$e'(t) \approx e(t-\Delta t)\exp(\lambda\Delta t). \tag{17}$$

Now let us analyze the other part e''. According to Eqs. (12) and (15), one obtains

$$\mathbf{y}'(t) \approx \mathbf{y}'(t-\Delta t) + \mathbf{F}[\mathbf{y}'(t-\Delta t)]\Delta t, \quad \mathbf{x}(t) \approx \mathbf{x}(t-\Delta t) + \{\mathbf{F}[\mathbf{x}(t-\Delta t)] - sk\mathbf{H}[\mathbf{x}(t-\Delta t)] + sk\mathbf{E}\}\Delta t.$$

By considering $\mathbf{y}'(t-\Delta t) = \mathbf{x}(t-\Delta t)$, one gets

$$\mathbf{y}'(t) - \mathbf{x}(t) \approx s\{\mathbf{H}[\mathbf{x}(t-\Delta t)] - \mathbf{E}\}k\Delta t.$$

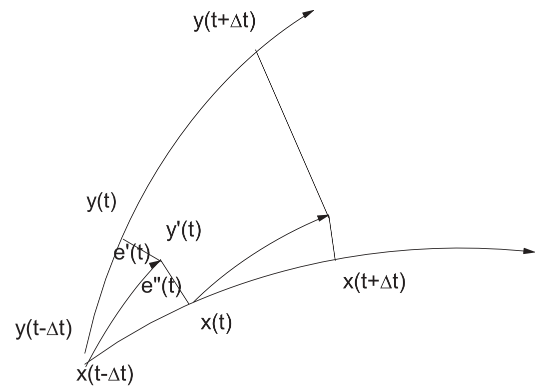

FIG. 7. A schematic plot of the trajectory analysis. Trajectories are plotted to illustrate e. $\mathbf{x}(t)$ is governed by Eq. (15). The dynamics of both $\mathbf{y}(t)$ and $\mathbf{y}'(t)$ are governed by Eq. (12), but $\mathbf{y}'(t)$ begins from the initial value $\mathbf{y}'(t-\Delta t) = \mathbf{x}(t-\Delta t)$, where Δt is a short time interval. $e'(t)$ and $e''(t)$ are Euclidean distances in m dimension.

Note that $\psi(t) = \|s\{\mathbf{H}[\mathbf{x}(t-\Delta t)] - \mathbf{E}\}\|_2$; the function $\psi(t) \geqslant 0$ is determined by the node and the sample. We have

$$e''(t) \approx \psi(t)k\Delta t. \tag{18}$$

By substituting Eqs. (17) and (18) into $e(t)$, we get $\Delta e = e(t) - e(t-\Delta t) = e(t-\Delta t)[\exp(\lambda\Delta t) - 1] + \psi(t)k\Delta t$. In the short time limit $\Delta t \to 0$, one has $\exp(\lambda\Delta t) - 1 \approx \lambda\Delta t$, and then $\dot{e} \approx e\lambda + \psi(t)k$. By integrating this equation and considering the initial condition $e(0) = 0$,

$$e(t) \approx \frac{\psi'(\psi)k}{\lambda}[\exp(\lambda t) - 1], \tag{19}$$

where $\psi'(\psi) > 0$ is a constant dependent on the node and the sample. It can be easily found that $e(t)$ grows as long as the node dynamics is chaotic, i.e., $\lambda > 0$. The exponential term on the right-hand side implies that the difference $e(t)$ increases exponentially with time. One can thus set $e(t_0) = \theta$ in each sampling, where $\theta > 0$ is a constant independent of the node and the sample. Under the assumption that the time series of the output functions $\{\mathbf{H}[\mathbf{x}_i(t)]\}$ have the same probability distribution P_c, we have $s'' \equiv \langle\psi'\rangle_{\text{ensemble}}$, which is independent of node i. Taking the average over the ensemble of Eq. (19), we have

$$\theta \approx \frac{s''k}{\lambda}[\langle\exp(\lambda t_0)\rangle_{\text{ensemble}} - 1].$$

Because of the first-order approximation $\langle\exp(\lambda t_0)\rangle_{\text{ensemble}} \approx \exp(\langle\lambda t_0\rangle_{\text{ensemble}}) = \exp(\lambda T_1)$, we have

$$\theta \approx \frac{s''k}{\lambda}[\exp(\lambda T_1) - 1]. \tag{20}$$

Here only the degree k and the MFPT T_1 are related to the node, and other parameters are all related to the node dynamics. The formula (20) indicates that the relation is closely related to the maximum Lyapunov exponent of the node dynamics. Moreover, the MFPT T_1 is shown to be a function of the node degree k. Therefore, Eq. (20) reveals the relationship between the degree and the MFPT, and the Lyapunov exponent λ.

Since $1/\lambda$ and T_1 imply the relaxation time of the orbit divergence in a chaotic system and the relaxation time of the divergence of the evolution of the coupled network from the uncoupled state, respectively, the exponential term in Eq. (20) reflects the competition between these two time scales. It is instructive to study the consequence of this interesting time-scale competition. Let us focus on two limiting cases. When $T_1 \gg 1/\lambda$, $\exp(\lambda T_1) - 1 \approx \exp(\lambda T_1)$. Equation (20) is simplified to

$$T_1 \approx \frac{1}{\lambda}[-\ln(s''k) + \ln(\lambda\theta)] = -\frac{1}{\lambda}\ln(s'k), \tag{21}$$

where $s' = s''/\lambda\theta$. This logarithmic dependence conforms to the results of Lorenz and Chua's circuit networks. When $T_1 \ll 1/\lambda$, $\exp(\lambda T_1) - 1 \approx \lambda T_1$. Equation (20) is simplified to

$$T_1 \approx \frac{\theta}{s''k}.$$

This relation conforms to the results of Rössler networks given in Fig. 5(b). Of course the MFPT T_1 depends crucially on the choice of the threshold θ, and this gives rise to different time scales as compared to the time scale $1/\lambda$. For the case of two comparable time scales T_1 and $1/\lambda$, the relation (20) gives a quantitative description. It should be noted that the above formula is independent of the node dynamics and the coupling function. The function $T_1(k,\lambda,\theta)$ precisely figures out the relation between the node degree and the node dynamics.

The accuracy of the MFPT method can be analyzed as follows. We take the case of $T_1 \gg 1/\lambda$ as an example. Based on Eq. (21), $T_1 \propto \ln k$, then

$$\Delta T_1 \propto \ln[s'(k+1)] - \ln(s'k) = \ln[s'(1+1/k)].$$

This implies that nodes with lower degrees can be distinguished more clearly than those with higher degrees. However, because hubs in BA networks are always the minority, and Δk's among hubs are big, the determination of the hubs can still be done effectively by using the MFPT approach. An advantage of the MFPT method is that its accuracy can be improved practically by taking more samples according to the large-number law in probability theory, since T_1 is an averaged quantity. When we take 30 000 samples (as shown in Fig. 4), the distribution of T_1 with the same degree is narrower than that by taking 1000 samples, as given in Fig. 5(a).

It may be too strict to require oscillators to be identical and to start evolution from the identical initial states. In fact, these requirements can be effectively weakened in practice. According to ergodic theory, for any initial state $\mathbf{x}_i(0)$ in the attractor, there always exists a small difference $\delta > 0$, so that one can get an isolated trajectory $\{\mathbf{y}_i(t)\}$ starting from an initial state $\mathbf{y}_i(0)$ that satisfies the condition $\|\mathbf{y}_i(0) - \mathbf{x}_i(0)\| < \delta$ only if the time series $\{\mathbf{y}_i(t)\}$ is long enough. However, the small difference δ may affect the MFPT. In the above analysis, $\dot{e} = \lambda e + \psi k$. If $\delta\lambda \gg \psi k$, λe is the main part in $\dot{e}$ from the beginning, and the function of k is weak. Therefore, the estimation of small degrees will be affected by the difference δ more easily than larger ones (see Fig. 8).

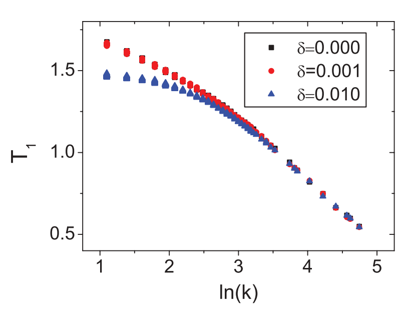

FIG. 8. (Color online) The estimation results from $\mathbf{e}(0) < \delta$. The simulation model is the same as that in Fig. 4 ($s = 0.001$).

V. CONCLUSIONS

In summary, we studied the relationship between dynamics and network topology with the goal of obtaining a dynamic estimation of the topological properties of a network. The bridge between these two different aspects of a system lies in the couplings among different nodes. In the absence of synchronies among nodes, one can uncover information on the links by analyzing the dynamical evolutions or the output time series. In terms of a probabilistic analysis of the coupling terms in the equations of motion, we reveal that a universal correspondence between the coupling term and the degrees of the nodes can be well established. This correspondence is independent of the specific forms of the node dynamics and the form of the output function. Furthermore, we proposed a method to estimate the degrees (in-degree) by using the MFPT. The relationship between the degree and the MFPT is investigated both numerically and theoretically. Our numerical explorations reveal that the MFPT exhibits a logarithmic dependence on the degree of the network for node dynamics described by Lorenz and Chua's oscillators, while an exponential dependence can be found for Rössler nodes. This can be well understood by using theoretical analysis. It is revealed that the MFPT is closely related to the largest Lyapunov exponent describing the deviation of the network dynamics from the isolated-node dynamics. We derived analytically the relation between the MFPT and the degree, where the largest Lyapunov exponent of an isolated-node oscillator λ plays an important role. Because $1/\lambda$ and T_1 imply the relaxation time of the divergence in a chaotic oscillator and the relaxation time of the divergence of coupled networks from the uncoupled state, the relationship between the MFPT and the degree reflects the competition between these two time scales. Our numerical studies of different node dynamics is found to fall into different limiting cases $\lambda T_1 \ll 1$ and $\lambda T_1 \gg 1$, respectively.

The MFPT method is efficient in judging the relative values of the node degrees and the degree distribution in a network. The computational cost of this approach is proportional to N, the number of nodes in a network. This method is applicable for chaotic node dynamics, and is independent of the specific forms of the node dynamics and the output functions. Moreover, we can naturally extend this approach to the cases in which explicit knowledge of node dynamics and coupling forms is absent. The proposed method greatly improves the practical topology-identification efficiency because in most circumstances it is difficult to obtain information on the node dynamics and the linking forms. It should be noted that even though the proposed method can be generalized to cases in which there is no explicit knowledge of the dynamics, the largest Lyapunov exponent of the system should be known for estimating the degree distribution. While if we fix T_1, the mean value of θ is just proportional to the node degree k according to Eq. (20). In this case the degree distrabution can be estimated by θ directly, so the value of the largest Lyapunov exponent is not required.

Up to now, most of the existing methods for estimating the degree of nodes in complex networks have been focused on in-degree estimations [12,13], and they are all based on the dynamical equations of motion proposed by Eq. (1). It is a challenge to estimate the out-degrees and other important statistical topology parameters on complex networks. Another interesting issue is the estimation of modules and clusters in a network by using our MFPT approach. The present approach is based on the requirement of irregular network dynamics because the concept of MFPT is proposed in a statistical sense. Discovering how to identify the network topology with periodic or excitable node dynamics remains an open and challenging issue.

ACKNOWLEDGMENTS

We wish to thank Professor Changsong Zhou at Hong Kong Baptist University for his suggestions on the generalized method. This work has been financially supported by the National Natural Science Foundation of China (Grants No. 10875011 and No. 11075016), the 973 project (2007CB814805), the Fundamental Research Funds for the Central Universities, and the Foundation for Doctoral Training from MOE.

[1] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).
[2] A.-L. Barabsi and R. Albert, Science **286**, 509 (1999).
[3] J. F. Donges, Y. Zou, N. Marwan, and J. Kurths, Eur. Phys. J. - Special Topics **174**, 157 (2009).
[4] J. F. Donges, Y. Zou, N. Marwan, and J. Kurths, Europhys. Lett. **87**, 48007 (2009).
[5] N. Wessel, A. Suhrbier, M. Riedl, N. Marwan, H. Malberg, G. Bretthauer, T. Penzel, and J. Kurths, Europhys. Lett. **87**, 10004 (2009).
[6] A. Arenas, A. Díaz-Guilera, and C. J. Pérez-Vicente, Phys. Rev. Lett. **96**, 114102 (2006).
[7] M. C. Romano, M. Thiel, J. Kurths, and C. Grebogi, Phys. Rev. E **76**, 036211 (2007).
[8] D. Yu, M. Righero, and L. Kocarev, Phys. Rev. Lett. **97**, 188701 (2006).
[9] M. Timme, Phys. Rev. Lett. **98**, 224101 (2007).
[10] S. G. Shandilya and M. Timme, New J. Phys. **13**, 013004 (2011).
[11] J. Ren, W.-X. Wang, B. Li, and Y.-C. Lai, Phys. Rev. Lett. **104**, 058701 (2010).
[12] S.-L. Bu and I.-M. Jiang, Europhys. Lett. **82**, 68001 (2008).
[13] Y. Shen, Z. Hou, and H. Xin, Chaos **20**, 013110 (2010).
[14] X. Liang, Z. Liu, and B. Li, Phys. Rev. E **80**, 046102 (2009).
[15] H.Risken, *The Fokker-Planck Equation* (Springer-Verlag, Berlin, 1984).
[16] Z. Zheng, G. Hu, and B. Hu, Phys. Rev. Lett. **81**, 5318 (1998).
[17] M. Zhan, G. Hu, Y. Zhang, and D. He, Phys. Rev. Lett. **86**, 1510 (2001).
[18] C. Zhou and J. Kurths, Phys. Rev. Lett. **96**, 164102 (2006).
[19] L. Chen, J.-a. Lu, and C. Tse, IEEE Trans. Circ. Syst. II **56**, 310 (2009).
[20] L. Huang, Q. Chen, Y.-C. Lai, and L. M. Pecora, Phys. Rev. E **80**, 036204 (2009).
[21] P. Erdös and A. Rényi, Publ. Math. Debrecen **6**, 290 (1959).
[22] T. Kapitaniak, *Controlling Chaos* (Academic, London, 1996).