Weak signal transmission in complex networks and its application in detecting connectivity

Xiaoming Liang,[1] Zonghua Liu,[1,*] and Baowen Li[2,3]
[1]*Institute of Theoretical Physics and Department of Physics, East China Normal University, Shanghai 200062, China*
[2]*Department of Physics and Centre for Computational Science and Engineering, National University of Singapore, Singapore 117546, Republic of Singapore*
[3]*NUS Graduate School for Integrative Sciences and Engineering, Singapore 117456, Republic of Singapore*
(Received 15 October 2008; revised manuscript received 26 March 2009; published 6 October 2009)

We present a network model of coupled oscillators to study how a weak signal is transmitted in complex networks. Through both theoretical analysis and numerical simulations, we find that the response of other nodes to the weak signal decays exponentially with their topological distance to the signal source and the coupling strength between two neighboring nodes can be figured out by the responses. This finding can be conveniently used to detect the topology of unknown network, such as the degree distribution, clustering coefficient and community structure, etc., by repeatedly choosing different nodes as the signal source. Through four typical networks, i.e., the regular one dimensional, small world, random, and scale-free networks, we show that the features of network can be approximately given by investigating many fewer nodes than the network size, thus our approach to detect the topology of unknown network may be efficient in practical situations with large network size.

DOI: 10.1103/PhysRevE.80.046102 PACS number(s): 89.75.Fb, 05.45.Xt

I. INTRODUCTION

The complex networks have been intensively studied in the past decade. It has been found that the dynamics on complex network can be significantly influenced by its topological structure, such as synchronization, epidemic spreading, and packet delivering, etc. [1–3]. For most of the biological systems, the structure and dynamics may influence each other to form an optimal functional network, i.e., the final observed network is a consequence of their interactions. Assuming that the topology of the network is unknown, the dynamics might be used to refer the connectivity [4–8].

Revealing connectivity is very important in the networks of neurons, interacting proteins or genes, ecological foodwebs, and even the functional network of brain, etc., in which the important aspects of network structure are largely unknown. To know their connectivity is of great help to understand their behaviors. For example, the prediction of epileptic seizure is still an unsolved problem although it has been studied for a long time. Finding out how the connectivity of functional brain network changes between the normal and abnormal function is the key to the prediction of seizure [9,10]. The problem of revealing connectivity has been intensively studied recently and several approaches have been presented [4–8]. For example, Makarov *et al.* proposed a method to identify the effective connectivity of neural network by using extracellular spike recordings [4]. Yu *et al.* estimated the topology of networks by designing control signals at each node [5]. Timme present an approach to infer the connectivity by introducing constant external driving [7]. All of these approaches need the global dynamics/information among all the pairs of nodes to figure out the connectivity, which cost a heavy calculation when the network size is large. Therefore, an interesting question is whether it is possible to give a simple method to reveal the links of a node directly, i.e., tell the links of a source node directly from the response of other nodes. This question is especially important for the networks with large size where the global structure of network is unknown and attention is usually paid to local areas or subnetworks.

On the other hand, signal transmission is of interesting from both a fundamental and a clinical perspective, such as in biological systems, and has been well studied in nonlinear science. It is found that noise can sometimes enhance the signal propagation [11–16]. However, previous works are limited to one-dimensional (1D) or two-dimensional (2D) regular lattices. Considering that many realistic systems can be simplified as networks, it is interesting to know how signal is transmitted in complex networks. This problem has been addressed recently. For examples, Acebron *et al.* studied the situation of scale-free (SF) network where all the nodes receive the same external signal and found that the amplitude amplification at the hub shows a resonance on the coupling strength [17]. Batista *et al.* investigated the spiking-bursting activity of neurons in a SF network with an external signal and found the appearance of frequency locking between the bursting and driving phases [18]. And one of us discussed the case of SF network with both signal and noise and found a double-resonance phenomenon [19,20]. What has been lacking is how a weak signal is transmitted in complex networks.

Weak signals appear at many situations. For example, in stochastic resonance, we usually consider the weak signal in a noisy environment and aim to find an optimal noise strength for the maximum signal-to-noise ratio [14,21–23]. In the visual and auditory systems of animals, there is a high sensitivity to weak external signals [17,19]. Signal may be produced by different ways. In biological systems, signal is often produced in a local area and then spread to other parts, such as pacemaker in human heart and epileptic foci in brain, etc. [24,25].

*zhliu@phy.ecnu.edu.cn

In this paper, we will present a simple network model of coupled oscillators to show how a weak signal propagates through the network. As an explicit example, we consider networks of coupled double-well systems, a paradigmatic model that has been successfully used to understand the stochastic resonance of signal detection [11–16] and signal transmission [17,19]. The double-well system has two wells and thus, has two distinct behaviors. One is a local oscillation limited in one of the two wells, another is a jumping behavior between the two wells. When the signal is not large enough to induce the jumping behavior, an external noise may help to induce the jumping and make the jumping to follow the frequency of signal at an optimal noise strength. This characteristic jumping behavior in the double-well system can be used to model the two-state behaviors in biological systems, such as fireflies, cicadas, crickets, tree-frogs, and even neurons [26,27]. The behaviors of these systems can be simplified as two states, i.e., quiescence and firing. It has been pointed out that the firing transfers the information and the residence-time distribution of a double-well system can exhibit the main features of the interspike interval histograms of firing neurons.

The coupled oscillators are usually used to study the synchronization by increasing coupling strength. Little attention has been paid to the desynchronization at weak coupling strength. We here consider the case of weak coupling in our model and aim to show the usefulness of desynchronization. In detail, we randomly choose a node from the network to add signal and investigate its subsequent influence on other nodes. Very interesting we find that if both the signal and coupling are weak, the signal transmission will be mainly limited to its nearest neighbors and thus, provides an efficient approach to detect the connectivity by directly measuring the response of other nodes to the source signal. We show both theoretically and numerically that the transmission of weak signal in networked double-well system decays exponentially with the topological distance to the source. We then use this finding to reveal the topology of network, such as the degree distribution, clustering coefficient and community structure, etc., by repeatedly choosing different nodes as the signal sources for "pinging" of the network. We find that the information on the network topology can be figured out by a number of sources much smaller than the network size. This approach is thus useful in practical situations where the network size N is very large.

The paper is organized as follows. In Sec. II, we present the model and show both theoretically and numerically that the exponentially decay of signal transmission and the detection of coupling strength. Then in Sec. III, we show how the approach can be used to detect the unknown networks, such as the degree distribution, clustering coefficient and community structure, etc. Finally we summarize our results in Sec. IV.

II. MODEL

Consider a network with hidden links among nodes, i.e., the network connectivity is unknown. In this case, all the information about the connectivity must be gathered from

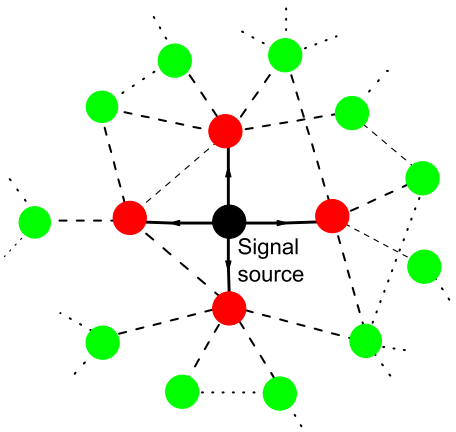

FIG. 1. (Color online) Schematic illustration of how a signal is transmitted in an unknown network. The solid lines represent the detected links by the response of the nearest neighbors, the dashed lines denote the hidden links connected to the nearest neighbors, and the dotted lines denote the hidden links connected to the nearest neighbors' neighbors, and so on.

"sensors" placed at each node. A key problem is what kind of oscillators can be used as the sensors and how the sensors tell us the hidden links. We here present an effective approach to solve this problem. Our idea to figure out the hidden links can be illustrated by the schematic Fig. 1 where the "dashed" and "dotted" lines represent the hidden links. We randomly choose a node as the source node to add a signal. Then we measure the "response" of other nodes and from that to figure out the hidden links. The signal will propagate from the source node to its nearest neighbors (see the "arrows" in the center of Fig. 1), and then to the nearest neighbors' neighbors, and so on. Suppose the response of nodes decays exponentially with their distance from the signal source. The response at the nearest neighbors will be much larger than that at the other nodes, thus, we can figure out all the links connected to the source node and then replace them with the "solid" lines, see Fig. 1. After that, we choose another node as the source node and do the same steps to figure out its links. In this way, we can figure out all the links of the network by repeatedly choosing different nodes as the signal sources.

As an example to implement the above steps, we here consider a network model of N nodes where each node is occupied by a double-well oscillator. We randomly choose one node to which a weak signal $A \sin \omega t$ is added. The evolution equation at each node can be described as follows:

$$\dot{x}_i = x_i - x_i^3 + \delta(i - i')A \sin \omega t + \varepsilon \sum_{j=1}^{N} h_{ij}(x_j - x_i), \qquad (1)$$

where $i=1,2,\ldots,N$, δ is the Kronecker delta function, x_i is the dynamical variable of the node i, and ε is the coupling strength. The topology of network connections is determined by the adjacency matrix $H=(h_{ij})$: $h_{ij}=1$ if the node j is connected to the node i, and $h_{ij}=0$ otherwise. The external signal $A \sin \omega t$ is added to the node i'. We here consider the situation with weak signal and weak coupling strength to mimic

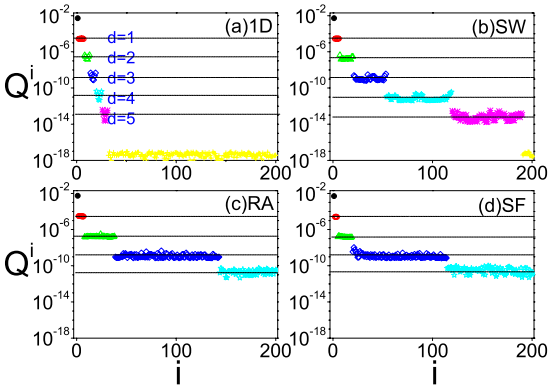

FIG. 2. (Color online) Q_i versus i for one pinging with $N=201$ and $\varepsilon=0.01$ where the i is ordered by the distance to the source node and the "red circles," "green triangles," "blue diamonds," "cyan stars," and "magenta asterisks" represent the distance $d=1$–5 to the "black" source node and the "yellow pluses" denotes the distance $d>5$. (a) Regular 1D network ($p=0.0$), (b) small-world network ($p=0.1$), (c) random network ($p=1.0$), and (d) scale-free network.

the slow spreading of signal in some biological systems where synchronization will not be resulted in.

The added signal at the node i' will be transmitted to its neighbors and then to the neighbors' neighbors and so on. After the transient process, the system will reach a stationary state. For measuring the transmission of signal, we calculate the response of node i at the reference frequency [12,13]

$$Q^i = \sqrt{[Q^i_{\sin}]^2 + [Q^i_{\cos}]^2} \tag{2}$$

with $Q^i_{\sin} = \frac{1}{T}\int_0^T [x_i(t) - \bar{x}_i]\sin\,\omega t dt$, $Q^i_{\cos} = \frac{1}{T}\int_0^T [x_i(t) - \bar{x}_i]\cos\,\omega t dt$, and $\bar{x}_i = \frac{1}{T}\int_0^T x_i(t)dt$. In numerical simulations, we fix $\omega=0.05$ and $T=2\pi/\omega$. We let $A=0.02$ so that the signal is weak enough to make the source node oscillate only around one of its two equilibriums. As examples, we consider four typical networks, i.e., the regular 1D, small world (SW), random (RA), and SF networks. The first three are constructed by the Watts and Strogatz's algorithm [28] where the 1D network has degree $k=6$ and the SW and RA networks are formed by rewiring each link of 1D network with probability $p=0.1$ and 1.0, respectively. The SF network is constructed by the Barabasi and Albert's (BA) algorithm [29] with average degree $\langle k\rangle=6$ and degree distribution $P(k) \sim k^{-3}$. Figure 2 shows the response Q^i versus i for $N=201$ and $\varepsilon=0.01$ where the horizontal axis i is ordered by the distance to the source node and (a) to (d) represent the cases of 1D, SW, RA, and SF networks, respectively. For clearness, we let d be the distance of the node i to the signal source i' and use the "red circles," "green triangles," "blue diamonds," "cyan stars," and "magenta asterisks" to represent the nodes located at $d=1\sim5$, respectively, and use the "yellow pluses" to represent the nodes at $d>5$. From Fig. 2 it is easy to see that Q^i in all the cases are distributed at different layers with the source at the top, its nearest neighbors at the first layer, and so on, see the lines with $d=1,2,\ldots$ in Fig. 2(a). Replace Q^i by $Q^i(d)$. Obviously, all the $Q^i(d)$ in Figs. 2(a)–2(d) decay exponentially with d.

To understand the mechanism why $Q^i(d)$ decays exponentially with d, we transfer the network to a topological structure with different layers and use X_d to denote the dynamical variable of a node at the layer d. Noticing that $Q^i(d) \gg Q^j(d+1)$ and $Q^i(d) \approx Q^j(d)$ in Figs. 2(a)–2(d), we assume that the response to the signal at the layer-d is mainly coming from the layer-$(d-1)$. Thus, we ignore the influence from the layers-d and -$(d+1)$ and simplify the Eq. (1) as follows:

$$\dot{X}_0 = (1 - k_0\varepsilon)X_0 - X_0^3 + A\,\sin\,\omega t,$$

$$\dot{X}_d = (1 - k_d\varepsilon)X_d - X_d^3 + \alpha_d\varepsilon X_{d-1}, \tag{3}$$

where k_0 denotes the degree of the source node, k_d the degree of a node at the layer-d, α_d the average links from the layer-$(d-1)$, and $d=1,2,\ldots$. Obviously, we have $\alpha_1=1$ and $\alpha_d \geq 1$ for $d \geq 2$.

When both A and ε in Eq. (3) are weak enough compared with the potential barrier of the double-well oscillator, $X_d(t)$ will be conveniently linearized around one of the potential minima. Then, the X_0 in Eq. (3) can be solved and asymptotically for long-time yields

$$X_0(t) \sim X_0(0) + \frac{A}{\sqrt{(2+k_0\varepsilon)^2 + \omega^2}}\sin(\omega t - \varphi_0), \tag{4}$$

where $\tan(\varphi_0) = \omega/(2+k_0\varepsilon)$ and $X_0(0) = \pm 1$ depending on the initial condition. Inserting this solution into the X_1 in Eq. (3) we obtain

$$X_1(t) \sim X_1(0) + \frac{A\varepsilon}{\sqrt{\prod_{i=0}^{1}[(2+k_i\varepsilon)^2 + \omega^2]}}\sin(\omega t - \varphi_0 - \varphi_1), \tag{5}$$

where $\tan(\varphi_1) = \omega/(2+k_1\varepsilon)$. Similarly, we have

$$X_d(t) \sim X_d(0) + \frac{A\prod_{i=1}^{d}\alpha_d\varepsilon}{\sqrt{\prod_{i=0}^{d}[(2+k_i\varepsilon)^2 + \omega^2]}}\sin\left(\omega t - \sum_{i=0}^{d}\varphi_i\right), \tag{6}$$

where $\tan(\varphi_i) = \omega/(2+k_i\varepsilon)$. Substituting Eqs. (4)–(6) into Eq. (2) we obtain

$$Q(d) = \frac{A\prod_{i=1}^{d}\alpha_d\varepsilon}{2\sqrt{\prod_{i=0}^{d}[(2+k_i\varepsilon)^2 + \omega^2]}}. \tag{7}$$

If we approximately consider $\alpha_d=\alpha$ and $k_i=K$, we have $\ln Q(d) = a(K) - b(K)d$ with $a(K) = \ln\frac{A}{2\sqrt{(2+K\varepsilon)^2+\omega^2}}$ and $b(K) = -\ln\frac{\alpha\varepsilon}{\sqrt{(2+K\varepsilon)^2+\omega^2}}$. That is, $Q(d)$ decays exponentially with the increase of d, confirming what we have observed in Figs. 2(a)–2(d).

From Eq. (7) we see that for the case of identical coupling in all links, the coupling strength can be evaluated through

$$\varepsilon = \frac{2k_1\beta^2 + \sqrt{4\beta^2 + \omega^2(1-\beta^2 k_1^2)}}{1-\beta^2 k_1^2}, \tag{8}$$

where $\beta = Q(1)/Q(0)$. For homogeneous network, we have $k_1 \approx k_0$ and thus, from Eq. (7) we obtain

$$\varepsilon = \frac{AQ(1)}{2Q(0)^2}. \tag{9}$$

While for the case of nonidentical coupling in different links, Eq. (3) can be approximately written as

$$\dot{X}_0^i = (1 - k_i\bar{\varepsilon}_i)X_0^i - (X_0^i)^3 + A\sin\omega t,$$

$$\dot{X}_1^j = (1 - k_j\bar{\varepsilon}_j)X_1^j - (X_1^j)^3 + \varepsilon_{i,j}X_0^i, \tag{10}$$

where X_0^i denotes the signal source, X_1^j one neighbor of node i, and $\bar{\varepsilon}_i$ and $\bar{\varepsilon}_j$ are the average coupling strength of the nodes i and j, respectively. Doing similar derivative with Eqs. (7) and (8) we obtain

$$Q^i(0) = \frac{A}{2\sqrt{(2 + k_i\bar{\varepsilon}_i)^2 + \omega^2}},$$

$$Q^j(1) = \frac{\varepsilon_{i,j}Q(0)}{\sqrt{(2 + k_j\bar{\varepsilon}_j)^2 + \omega^2}}. \tag{11}$$

When $k_j \approx k_i$, the coupling strength can be calculated approximately through

$$\varepsilon_{i,j} = \frac{AQ^j(1)}{2Q^i(0)^2}. \tag{12}$$

When k_j is largely different from k_i, such as the SF network, Eqs. (9) and (12) will not work again. However, considering that a SF network has only a few hubs, the degree differences among the general nodes, excluding the hubs, are not very large. We will use numerical simulations to confirm this argument. In the following numerical simulations, we do statistical average to get the needed quantity. That is, for a set of parameters we make 20 networks or realizations and for each network we make N times pinging. The measured quantities are averaged on ensemble. We consider the BA network with $N=201$ and calculate its histogram of degree difference Δk. Figure 3(a) shows the result. It is easy to see that the $P(\Delta k)$ with $\Delta k \leq 10$ takes most of the possibility. That is, for a randomly chosen source node, it is very possible that its k_0 and k_1 has $\Delta k \leq 10$, indicating the Eqs. (9) and (12) may be still useful. This point can be also supported by Eq. (11) for the case of weak coupling. In that situation, the small difference between k_i and k_j (i.e., $\Delta k \leq 10$) will not make a significant difference to the two denominators of Eq. (11), i.e., $|k_i\bar{\varepsilon}_i - k_j\bar{\varepsilon}_j| \ll 2$, indicating that Eq. (12) is still approximate correct. For checking the correctness of this argument, we consider the case of identical coupling and calculate the error $\Delta\varepsilon$ between the real ε and the estimated ε', i.e., $\Delta\varepsilon = \langle|\varepsilon' - \varepsilon|\rangle$. Figure 3(b) shows how the relative error $\Delta\varepsilon/\varepsilon$ changes with Δk, where the "circles" denote the case of $\varepsilon = 0.01$ and the "squares" the case of $\varepsilon = 0.02$. Obviously, $\Delta\varepsilon/\varepsilon$ is very small for most of Δk except the hubs, confirming our argument.

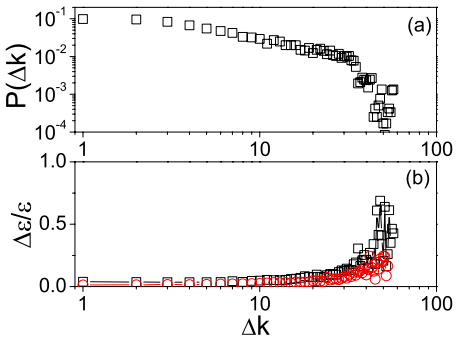

FIG. 3. (Color online) (a) $P(\Delta k)$ versus Δk and (b) $\Delta\varepsilon/\varepsilon$ versus Δk for BA network with $N=201$ and averaged over 20 realizations. In (b), the "circles" denote the case of $\varepsilon=0.01$ and the "squares" the case of $\varepsilon=0.02$.

We have also checked the other three typical networks. Figure 4 shows the results with error bars. Obviously, ε' is approximately equal to ε for the 1D, SW, and RA networks, and also for the SF network with $\varepsilon < 0.02$. Therefore, Eq. (9) is approximate correct for all the networks with weak coupling, i.e., $\varepsilon < 0.02$. Does this condition of weak coupling work for the case of nonidentical coupling in different links? For convenience, we here discuss the situation of symmetric coupling with $\varepsilon_{i,j} = \varepsilon_{j,i}$ and denote the real coupling strength as $\varepsilon_{i,j}$ and the predicted coupling strength from Eq. (12) as $\varepsilon'_{i,j}$. For one pinging with signal added at the node i, we can get k_i coupling strength $\varepsilon'_{i,j}$ from Eq. (12). After N times pinging with each node as the source node, respectively, we can get all the $\varepsilon'_{i,j}$. In numerical simulations, we let $\varepsilon_{i,j}$ be uniform randomly chosen from [0.005, 0.015] and calculate $\varepsilon'_{i,j}$ by $(\varepsilon'_{i,j} + \varepsilon'_{j,i})/2$. Figure 5 shows the relationship between the predicted $\varepsilon'_{i,j}$ by Eq. (12) and the real $\varepsilon_{i,j}$ where (a)–(d) represent the cases of 1D, SW, RA, and SF, respectively. Obviously, most of the pairs $\varepsilon'_{i,j} \sim \varepsilon_{i,j}$ are distributed around the diagonal line, indicating the consistence between the theoretical prediction Eq. (12) and the real couplings.

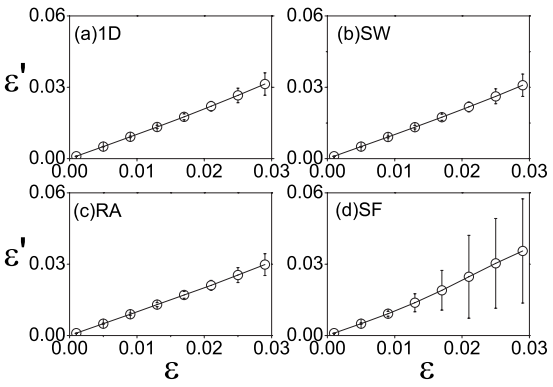

FIG. 4. Predicted ε' by Eq. (9) versus the real ε for $N=201$ and averaged over 20 realizations, where (a)–(d) represent the cases of 1D, SW, RA, and SF networks, respectively, and the diagonal line denotes the ideal case.

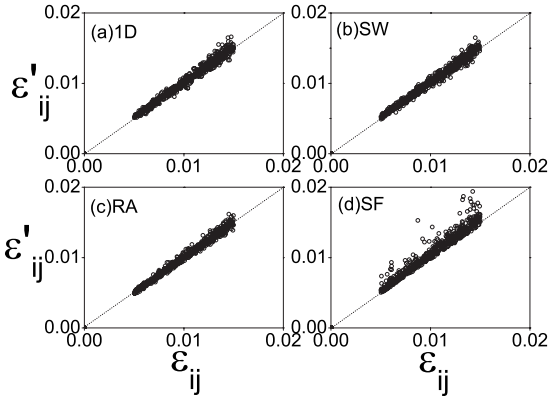

FIG. 5. Predicted $\varepsilon'_{i,j}$ by Eq. (12) versus the real $\varepsilon_{i,j}$ for $N=201$, where (a)–(d) represent the cases of 1D, SW, RA, and SF networks, respectively, and the diagonal line denotes the ideal case.

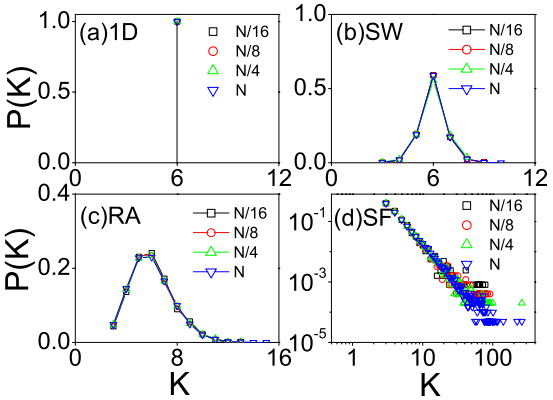

FIG. 6. (Color online) The measured degree distribution by pinged times $N_0=N/16$, $N/8$, $N/4$, and N with $N=1001$, $\varepsilon=0.01$ and (a) to (d) represent the 1D, SW, RA, and SF network, respectively.

III. DETECTION OF NETWORK STRUCTURE

When the network size N is very large, it is difficult to make N times pinging and thus in this situation how to obtain the information on the network topology, such as the degree distribution, clustering coefficient and community structure, etc., becomes extremely important. We find that the feature of exponential decay in the response $Q^i(d)$ may provide a great help in solving this problem. From Fig. 2, one can easily get some information on the degree distribution $P(k)$. Let $n(d)$ denote the number of nodes at the layer-d in Fig. 2. Obviously, $n(d)$ is a constant for $d=1\sim5$ in Fig. 2(a), increases approximate linearly with d in (b), increases with an approximate power-law of d in (c), and increases faster than power-law of d in (d). Thus, even one realization can tell us some primary information on the topology of the unknown network. To obtain more detailed information, we may use the network sampling approaches which include the "vertex sampling," "edge sampling," and "snowball-sampling," etc. [30–32].

To estimate the degree distribution, we use the vertex sampling to randomly choose N_0 nodes as the source nodes. For each chosen source node, we focus only on the first layer, i.e., $n(1)$. Noticing that $n(1)$ is in fact the degree of the source node k_0, the obtained $n(1)$ from the N_0 pinging will tell us the information of degree distribution. Figure 6 shows the histogram of $n(1)$ at pinged times $N_0=N/16$, $N/8$, $N/4$ and N, respectively. Obviously, the four cases in each panel of Fig. 6 have the similar shape, indicating that an approximate $P(k)$ can be measured by only one tenth pinging of the total nodes N.

The clustering coefficient represents the probability for two of one's neighbors to be neighbors also of each other. For an individual pinging, $n(1)$ nodes are all the neighbors of the source node. If the number of links, E, among the $n(1)$ nodes can be figured out, we will have

$$C_i = \frac{2E}{n(1)(n(1)-1)}. \tag{13}$$

An easy way to obtain E is that we let the $n(1)$ nodes be the signal source, respectively, and only measure the links among the $n(1)$ nodes, i.e., edge sampling. Then we randomly choose another source node and do the same procedure. After m_0 times pinging, we obtain

$$C(m_0) = \frac{1}{m_0}\sum_{i=1}^{m_0} C_i. \tag{14}$$

An interesting question is how fast $C(m_0)$ approaches the clustering coefficient C of the network. To answer this question, we use the SF network as an example to calculate $C(m_0)$. Figure 7(a) shows how $C(m_0)$ changes with m_0. From Fig. 7(a) it is easy to see that $C(m_0)$ converges very fast and approaches to C when $m_0<N/10$. As our SF network is constructed by the BA algorithm [29], it has a very small $C=0.034$. Considering that the unknown network may have a larger C, we here increase the value of C but let $P(k)$ remain unchanged by the rewiring approach given in [33,34]. Figures 7(b) and 7(c) show how $C(m_0)$ changes with m_0 for $C=0.25$ and 0.5, respectively. Obviously, $C(m_0)$ also converges very fast and approach to C in both Figs. 7(b) and

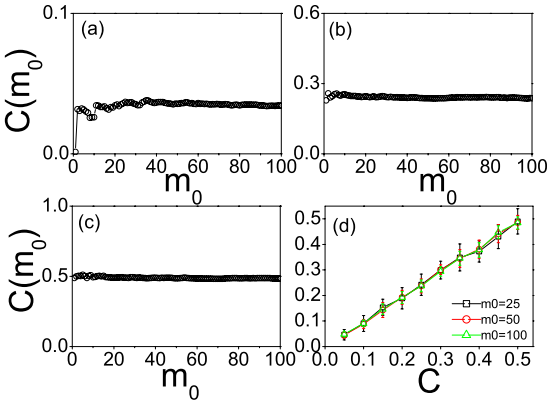

FIG. 7. (Color online) (a)–(c) The measured $C(m_0)$ versus m_0 for SF network with $C=0.034$ in (a), 0.25 in (b) and 0.5 in (c); (d) the measured $C(m_0)$ versus the real C of the network where the "circles," "triangles," and "stars" represent the cases of $m_0=25$, 50, and 100, respectively. The parameters are $N=1001$ and $\varepsilon=0.01$.

7(c). To compare the difference between $C(m_0)$ and the real C, we show $C(m_0)$ versus C in Fig. 7(d) for three cases of $m_0=25$, 50 and 100, respectively. It is clear that all the symbols in Fig. 7(d) are around the diagonal line, indicating that the clustering coefficient can be approximately figured out by a small number of pinging $m_0=\frac{25}{1000}N=\frac{1}{40}N$.

Community structures are prevalent in many real networks, such as social and biological networks, and is thus another feature of complex networks. Our approach can be also used to detect the existence of community structures in an unknown network. The characteristic feature of a community network is that it consists of groups and there is dense connections in a group and few connections between the groups [35,36]. Many algorithms have been proposed to detect the community structure of a network [36–38]. A basic assumption in these algorithms is that all the links or adjacency matrix are known. However, this condition is not always guaranteed. For example, for the situations considered in this paper, we do not know the connectivity and thus these algorithms cannot be directly used here.

Our idea to detect the community structures is through the snowball sampling and can be described as follows: If there is community structures in the unknown network, a random chosen source node must be in a specific group. Starting from this source node, we have shown that we can reveal its neighbors, i.e., the nodes at the first layer with $l=1$, by adding the weak signal in the source node. Then we let the nodes on the layer-1 be the source nodes, respectively, to reveal their links. By this way, one can get all the links of the local nodes around the starting node. At the same time we need to judge the "border" of the community. If the "border" is reached, the outgoing process of detecting community should be stopped; otherwise continue. Thus, a key element is how to judge which node is the "border" of the community. Fortunately, we find that the methods proposed by Refs. [38–44]. is just designed for our situation to determine the "border" of the community. We here would like to choose three typical approaches from them, i.e., the Clauset method in [39], the Luo, Wang, and Promislow (LWP) method in [40], and the Bagrow method in [42]. The common idea of the three approaches can be briefly summarized as follows: Start with a node i as the community $\mathcal{C}$ and detect its k_i neighbors as $\mathcal{B}$. At each subsequent step, one or more nodes from $\mathcal{B}$ are chosen and agglomerated into $\mathcal{C}$, then $\mathcal{B}$ is updated to include any newly discovered nodes. The difference of the three approaches is how to maximize the local modularity. In the following, we will briefly introduce their main procedure, respectively.

The Clauset method considers the local modularity [39]

$$R=\frac{\sum_{ij} B_{ij}\delta(i,j)}{\sum_{ij} B_{ij}}, \tag{15}$$

where B_{ij} is the element of boundary-adjacency matrix with at least one node in the boundary/border of $\mathcal{C}_{border}$, $\delta(i,j)$ is 1 when either a pair of connected nodes $i\in\mathcal{C}_{border}$ and $j\in\mathcal{C}$ or vice versa, and is 0 otherwise. Each node in $\mathcal{B}$ that can be agglomerated into $\mathcal{C}$ will cause a change in R, ΔR. At each time step, we calculate ΔR and let the node with the largest ΔR be agglomerated. The modularity R will lie in the range $[0, 1]$ and its local maxima indicate good community separation.

The LWP method introduces the number of edges internal and external to $\mathcal{C}$ as M_{in} and M_{out}, respectively:

$$M_{in}=\sum_{ij} S_{ij}\delta(i,j),$$

$$M_{out}=\sum_{ij} S_{ij}\lambda(i,j), \tag{16}$$

where S_{ij} is the element of the matrix with at least one node in $\mathcal{C}$, $\delta(i,j)$ is 1 if both node i and node j are in $\mathcal{C}$ and 0 otherwise, and $\lambda(i,j)$ is 1 only one of node i and node j belongs to $\mathcal{C}$ and 0 otherwise. The local modularity is defined as

$$M=\frac{M_{in}}{M_{out}}. \tag{17}$$

Its agglomerating process is a little complicated, see Ref. [40] for details. The algorithm will be repeated until no node is added to $\mathcal{C}$. If the M of the subnetwork $\mathcal{C}$ is large than 1 and $\mathcal{C}$ contains the started source node, the community is found. Otherwise, no community will be found for the started source node.

The Bagrow method can be stated as follows [42]. We first calculate the "outwardness" $\Omega_v(\mathcal{C})$ of node $v\in\mathcal{B}$ from community $\mathcal{C}$:

$$\Omega_v(\mathcal{C})=\frac{1}{k_v}\sum_{i\in n(v)}([i\notin\mathcal{C}]-[i\in\mathcal{C}])=\frac{1}{k_v}(k_v^{out}-k_v^{in}), \tag{18}$$

where $n(v)$ are the neighbors of v. We agglomerate the node with the smallest Ω at each step and update the community and its boundary. Then we calculate M_{out} and track M_{out} during agglomeration. The community has been fully agglomerated when M_{out} has a local minimum.

We here use the above three approaches to detect the community from the links revealed by our approach. One way that has been employed to test sensitivity in many cases is to see how well a particular method performs when applied to *ad hoc* networks with a well known, fixed community structure [39,41–44]. We here also use the *ad hoc* networks as an example to check our method. The *ad hoc* network consists of 128 nodes divided into four equally sized communities. Each node has degree $k=k_{in}+k_{out}=16$, where k_{out} is the links of a node outside its community. As its community $\mathcal{G}$ is known, we can calculate the normalized mutual information (NMI) [42,43] to compare the found community $\mathcal{C}$ and the real $\mathcal{G}$:

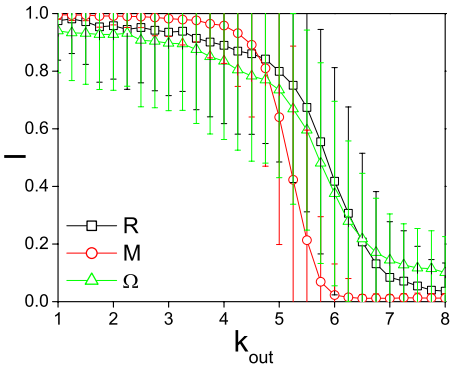

FIG. 8. (Color online) NMI versus k_{out} for three local methods for the *ad hoc* network with $N=128$, $\varepsilon=0.01$, and average degree $k=k_{in}+k_{out}=16$. R denotes the Clauset method, M the LWP method, and Ω the Bagrow method.

$$I(\mathcal{G},\mathcal{C}) = \frac{-2\sum_{ij} X_{ij} \ln\left(\frac{X_{ij}N}{X_{i.}X_{.j}}\right)}{\sum_{i} X_{i.} \ln\left(\frac{X_{i.}}{N}\right) + \sum_{j} X_{.j} \ln\left(\frac{X_{.j}}{N}\right)}, \qquad (19)$$

where X is a 2×2 matrix with X_{ij} being the number of nodes from real community i that detected in found community j, $X_{.j}=X_{1j}+X_{2j}$, and $X_{i.}=X_{i1}+X_{i2}$. Figure 8 shows the relationship between I and k_{out}. Notice that $I=1$ means a single community has been discovered while $I=0$ corresponds to the case where the algorithm's partition provides no additional information about the true partition. From Fig. 8 we see that all the three approaches give the similar result when $k_{out}\leq 5$, indicating the correctness of detected connectivity. We may need to point out that if the knowledge of the global structure of the network is known, the community can be found when k_{out} persists until $k_{out}=12$ [45,46].

Similarly, if we choose another source node which is not in the discovered community, we can find the second community. Doing this process continuously, we can find all the communities in principle. However, we have to point out that comparing with other approaches of finding all the communities [36–38], our approach is not so good. This disadvantage can be easily overcome as follows. For a finite size network, we can use our approach to reveal all the links and then use any of the known method [36–38] to figure out its communities. The advantage of our approach is to find the information on the topology of network, such as the degree distribution, clustering coefficient, and a particular community in a large network by local connections.

IV. DISCUSSIONS AND CONCLUSIONS

The results obtained in this paper are based on the condition of weak signal and weak coupling. Without this condition, the added signal may induce synchronization in the network and then our method cannot distinguish the links. That is, there is a critical coupling strength ε_c, which is related to the signal strength A. Our method works for $\varepsilon<\varepsilon_c$ in which we have $Q^i(d)\gg Q^j(d+1)$ and thus the nearest neighbors of the source node have much larger response than other bodes.

In conclusions, we have proposed a networked model of double-well systems to study the transmission of a weak signal in complex networks through weak coupling and found both numerically and theoretically that the response of the neighboring nodes decays exponentially and the coupling strengths can be measured. One advantage of exponential decay of response is that the nearest neighbors of the source can be easily distinguished from others. We have used this feature to show that the identification of network connectivity, such as the degree distribution, clustering coefficient and community structure, etc., can be implemented by a number of sources much smaller than the network size. This approach may be useful in the detection of connectivity for the case of unknown network with large size where it is difficult to let every node be the source for one time.

ACKNOWLEDGMENTS

This work was supported by the NNSF of China under Grants No. 10775052 and No. 10635040, and by National Basic Research Program of China (973 Program) under Grant No. 2007CB814800.

[1] R. Albert and A.-L. Barabasi, Rev. Mod. Phys. **74**, 47 (2002).
[2] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Phys. Rep. **424**, 175 (2006).
[3] Z. Liu, W. Ma, H. Zhang, Y. Sun, and P. M. Hui, Physica A **370**, 843 (2006); M. Tang, Z. Liu, and J. Zhou, Phys. Rev. E **74**, 036101 (2006); Z. Liu, Y.-C. Lai, and N. Ye, *ibid.* **67**, 031911 (2003).
[4] V. A. Makarov, F. Panetsos, and O. d. Feo, J. Neurosci. Methods **144**, 265 (2005).
[5] D. Yu, M. Righero, and L. Kocarev, Phys. Rev. Lett. **97**, 188701 (2006).
[6] A. Arenas, A. Diaz-Guilera, and C. J. Perez-Vicente, Phys. Rev. Lett. **96**, 114102 (2006).
[7] M. Timme, Phys. Rev. Lett. **98**, 224101 (2007).
[8] D. Napoletani and T. D. Sauer, Phys. Rev. E **77**, 026103 (2008).
[9] D. E. Lerner, Physica D **97**, 563 (1996).
[10] Z. Liu, B. Hu, and L. D. Iasemidis, Europhys. Lett. **71**, 200 (2005).
[11] J. F. Lindner, S. Chandramouli, A. R. Bulsara, M. Löcher, and W. L. Ditto, Phys. Rev. Lett. **81**, 5048 (1998).
[12] Y. Zhang, G. Hu, and L. Gammaitoni, Phys. Rev. E **58**, 2952 (1998).
[13] A. A. Zaikin, J. Garcia-Ojalvo, L. Schimansky-Geier, and J. Kurths, Phys. Rev. Lett. **88**, 010601 (2001).
[14] Z. Liu, Y.-C. Lai, and A. Nachman, Phys. Lett. A **297**, 75 (2002); Int. J. Bifurcation Chaos Appl. Sci. Eng. **14**, 1655 (2004).

[15] D. H. Zanette, Europhys. Lett. **68**, 356 (2004); Eur. Phys. J. B **43**, 97 (2005).
[16] P. Kaluza, M. Ipsen, M. Vingron, and A. S. Mikhailov, Phys. Rev. E **75**, 015101(R) (2007).
[17] J. A. Acebrón, S. Lozano, and A. Arenas, Phys. Rev. Lett. **99**, 128701 (2007).
[18] C. A. S. Batista, A. M. Batista, J. A. C. de Pontes, R. L. Viana, and S. R. Lopes, Phys. Rev. E **76**, 016218 (2007).
[19] Z. Liu and T. Munakata, Phys. Rev. E **78**, 046111 (2008).
[20] F. Lu and Z. Liu, Chin. Phys. Lett. **26**, 040503 (2009).
[21] L. Gammaitoni, P. Hanggi, P. Jung, and F. Marchesoni, Rev. Mod. Phys. **70**, 223 (1998).
[22] K. Park, Y.-C. Lai, Z. Liu, and A. Nachman, Phys. Lett. A **326**, 391 (2004).
[23] Y.-C. Lai, Z. Liu, A. Nachman, and L. Zhu, Int. J. Bifurcation Chaos Appl. Sci. Eng. **14**, 3519 (2004).
[24] E. Ben-Jacob, I. Doron, T. Gazit, E. Rephaeli, O. Sagher, and V. L. Towle, Phys. Rev. E **76**, 051903 (2007).
[25] M. Perc, Phys. Rev. E **78**, 036105 (2008).
[26] Z. Liu and P. M. Hui, Physica A **383**, 714 (2007).
[27] A. Longtin, A. Bulsara, D. Pierson, and F. Moss, Biol. Cyber. **70**, 569 (1994).
[28] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).
[29] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[30] O. Frank, in *Models and Methods in Social Network Analysis*, edited by P. J. Carrington, J. Scott, and S. Wasserman (Cambridge University Press, New York, 2005).
[31] S. H. Lee, P.-J. Kim, and H. Jeong, Phys. Rev. E **73**, 016102 (2006).
[32] L. F. Costa and G. Travieso, Phys. Rev. E **75**, 016102 (2007).
[33] S. Maslov and K. Sneppen, Science **296**, 910 (2002).
[34] B. J. Kim, Phys. Rev. E **69**, 045101(R) (2004).
[35] Z. Liu and B. Hu, Europhys. Lett. **72**, 315 (2005).
[36] M. E. J. Newman, Phys. Rev. E **67**, 026126 (2003); M. Girvan and M. E. J. Newman, Proc. Natl. Acad. Sci. U.S.A. **99**, 7821 (2002); M. E. J. Newman and M. Girvan, Phys. Rev. E **69**, 026113 (2004).
[37] Y. Hu, M. Li, P. Zhang, Y. Fan, and Z. Di, Phys. Rev. E **78**, 016115 (2008).
[38] J. P. Bagrow and E. M. Bollt, Phys. Rev. E **72**, 046108 (2005).
[39] A. Clauset, Phys. Rev. E **72**, 026132 (2005).
[40] F. Luo, J. Z. Wang, and E. Promislow, *Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI'06)* (IEEE, Piscataway, NJ, 2006).
[41] B. Karrer, E. Levina, and M. E. J. Newman, Phys. Rev. E **77**, 046119 (2008).
[42] J. P. Bagrow, J. Stat. Mech.: Theor. Exp. (2008), P05001.
[43] L. Danon, A. Diaz-Guilera, J. Duch, and A. Arenas, J. Stat. Mech.: Theor. Exp. (2005), P09008.
[44] M. Meila, J. Multivariate Anal. **98**, 873 (2007).
[45] J. Duch and A. Arenas, Phys. Rev. E **72**, 027104 (2005).
[46] R. Guimera and L. A. N. Amaral, J. Stat. Mech.: Theor. Exp. (2005), P02001.