Identifying partial topology of complex dynamical networks via a pinning mechanism

Shuaibing Zhu, Jin Zhou, and Jun-an Lu

Citation: Chaos **28**, 043108 (2018); doi: 10.1063/1.5009946
View online: https://doi.org/10.1063/1.5009946
View Table of Contents: http://aip.scitation.org/toc/cha/28/4
Published by the American Institute of Physics

CHAOS **28**, 043108 (2018)

Identifying partial topology of complex dynamical networks via a pinning mechanism

Shuaibing Zhu, Jin Zhou,[a)] and Jun-an Lu
School of Mathematics and Statistics, Wuhan University, Wuhan 430072, China

(Received 20 October 2017; accepted 23 March 2018; published online 10 April 2018)

In this paper, we study the problem of identifying the partial topology of complex dynamical networks via a pinning mechanism. By using the network synchronization theory and the adaptive feedback controlling method, we propose a method which can greatly reduce the number of nodes and observers in the response network. Particularly, this method can also identify the whole topology of complex networks. A theorem is established rigorously, from which some corollaries are also derived in order to make our method more cost-effective. Several numerical examples are provided to verify the effectiveness of the proposed method. In the simulation, an approach is also given to avoid possible identification failure caused by inner synchronization of the drive network. *Published by AIP Publishing*. https://doi.org/10.1063/1.5009946

Apart from synchronization, topology identification of complex networks is also of great importance. Sometimes, we are interested in only a part of the whole topology of complex networks. For example, we may want to know the friends of several people in social networks, the coauthors of several academics in coauthorship networks and the citations of several papers related to our research in citation networks. Research on synchronization-based topology identification of complex dynamical networks has been thriving in the past decade. Most of the existing studies focused on identifying the topology of the whole network. There are two fundamental drawbacks for most of the existing identification methods: (1) they are not practical when dealing with large-scale networks and (2) they are not cost-effective when we only want to know a part of the whole topology. Therefore, it is very necessary to develop a cost-effective method for identifying the partial topology of complex networks.

I. INTRODUCTION

Complex networks feature the complexity of having lots of nodes and complex topology. Complex networks exist in many natural and man-made systems, such as the World Wide Web, electrical power grids and neural networks.[1–3] In the past two decades, the research on complex networks has been thriving, and fruitful results were obtained.[4–10] Apart from synchronization,[11–17] which is a key topic, topology identification has received great attention.[18–30]

Unraveling the topological structure of complex networks is of practical importance in the real world. For instance, most diseases are the result of the collapse of cellular processes together with interaction networks among the components of the genome, the proteome and the metabolome.[31] Exploring the biological network between diseases, such as protein-protein interactions of chronic diseases, might give us a more detailed understanding of disease comorbidity.[32] Moreover, for many real-world complex networks, we may be interested in only a part of their whole topology. For example, when there are several newcomers in a social network, we may want to know the relationship between these newcomers and their friends or the relationship among them. In networks of coauthorship among academics, we may want to know the coauthors of several academics. Similarly, in citation networks, we may be interested in the citations of several papers that are closely related to our research. Therefore, the problem of identifying the partial topology of complex networks is of theoretical and practical importance.

In the past decade, a lot of studies have been devoted to synchronization-based topology identification. Many network models, including weighted complex dynamical networks,[18] networks with coupling delay,[19] networks with unknown system parameters,[22] and networks with stochastic perturbations,[23] have been investigated. The control scheme varies from impulsive control[21,24] to intermittent control[25] as well as pinning control.[26] However, there are two fundamental drawbacks for most of the aforementioned approaches. On the one hand, the response network always consists of the same amount of nodes as the drive network. When the drive network is of large scale, which is the common case, we need to construct a response network with many nodes and observers to estimate the unknown coupling matrix. For example, for a drive network consisting of 1000 nodes, there should be 1000 nodes and 1000×1000 observers in the response network, which is too consuming to be practical. On the other hand, most of these approaches identify the whole topology of complex networks, which is not cost-effective when we only want to know a part of the whole topology. They only tell us how to identify the whole topology, but they do not tell us how to use fewer nodes and observers to identify partial topology. Simply reducing the number of nodes in the response network is not feasible, which we will further explain. Zhou[26] proposed an approach for identifying the partial topology of a coupled FitzHugh-Nagumo neurobiological network via a pinning mechanism. This approach requires that a part of the topology be known

[a)]Electronic mail: jzhou@whu.edu.cn

beforehand, and the response network consists of the same amount of nodes as the drive network, making its practical value limited. Therefore, it is very necessary to develop a cost-effective method for identifying the partial topology of complex networks, which need fewer nodes and does not require that a part of the topology be known.

By using the network synchronization theory and the adaptive feedback controlling method, this paper proposes a novel and cost-effective method for identifying the partial topology of complex dynamical networks via a pinning mechanism, i.e., there are fewer nodes in the response network than in the drive network. Our main contributions are as follows. First, the design of the response network is novel. Since the response network has fewer nodes, the traditional concept of outer synchronization is no longer applicable. Our idea is to synchronize the response network with a part of the drive network. Second, we give an approach to avoid possible identification failure caused by inner synchronization of the drive network, while few studies addressed this issue.

The rest of this paper is organized as follows: Section II gives some necessary preliminaries. Section III proposes a method for identifying the partial topology of complex dynamical networks. Section IV provides several examples to verify the effectiveness of the proposed method. Section V gives some concluding remarks.

II. PRELIMINARIES

We first introduce some necessary notations that will be used throughout this paper. $\mathbb{R}^n$ and $\mathbb{R}^{n\times m}$ denote the n-dimensional Euclidean space and the set of all the $n\times m$-dimensional real matrices, respectively, $\mathbb{R}_+$ is the set of positive real numbers, the superscript T denotes the transpose of a vector or matrix, $||\cdot||$ denotes the 2-norm of a vector or matrix, $\lambda_{\max}(\cdot)$ denotes the maximal eigenvalue of a symmetric matrix, $\lambda_{\min}(\cdot)$ denotes the minimal eigenvalue of a symmetric matrix, $\otimes$ represents the Kronecker product, and I_N represents the identity matrix of dimension N.

Consider a complex network having N systems of dimension n. The equations describing the dynamical evolution of the network are

$$\dot{x}_i(t) = f_i(t, x_i(t)) + \sum_{j=1}^{N} c_{ij}\Gamma x_j(t), \tag{1}$$

where $1\le i\le N, x_i\in\mathbb{R}^n$ is the state vector of the i-th node, f_i: $\mathbb{R}_+\times\mathbb{R}^n\to\mathbb{R}^n$ is a smooth nonlinear function, the dynamics of the i-th node is $\dot{x}_i = f_i(t, x_i)$, Γ is the inner coupling matrix, and $C=(c_{ij})_{N\times N}$ is the unknown weight configuration matrix. If there is a link from node i to node j $(j\neq i)$, then $c_{ij}>0$ and c_{ij} is the weight; otherwise, $c_{ij}=0$, and the diagonal elements of matrix C are defined as

$$c_{ii} = -\sum_{j=1, j\neq i}^{N} c_{ij}, \quad i = 1, 2, \ldots, N.$$

Note that the configuration matrix C need not be symmetric or irreducible and Γ also need not be symmetric, but the boundedness of the network should be ensured in this paper.

Hereafter, network (1) is referred to as the drive network. Although the whole topology C is unknown, we might be interested in only a part of C. Given the partial topology that we are interested in, the main goal is to identify it in a cost-effective way. Suppose we are interested in the couplings between l $(1<l<N)$ nodes and their neighbors. Without loss of generality, these nodes are numbered as $1, 2, \ldots, l$, respectively. Construct another complex dynamical network composed of l nodes, referred to as the response network, and described by

$$\dot{y}_i(t) = f_i(t, y_i(t)) + \sum_{j=1}^{l} \hat{c}_{ij}\Gamma y_j(t) + \sum_{j=l+1}^{N} \hat{c}_{ij}\Gamma x_j(t) + u_i, \tag{2}$$

where $1\le i\le l, y_i\in\mathbb{R}^n$ is the response state vector of the i-th node of the drive network, u_i is the control input to be designed, and $\hat{c}_{ij}$ is the estimation of the weight c_{ij}. Then $C_{l\times N} = (c_{ij})_{l\times N}$ is the partial topology to be identified by $\hat{C}_{l\times N} = (\hat{c}_{ij})_{l\times N}$. Sometimes, we would like to know the relationship among l nodes, then only $C_{l\times l} = (c_{ij})_{l\times l}$ is needed, which is included in $C_{l\times N}$.

Remark 1: Figure 1 shows how the response network connects with the drive network. It is noteworthy that the observers $\hat{c}_{ij}(j>l)$ are indispensable due to two reasons. First, we need to identify $c_{ij}(j>l)$. Second, even if only $C_{l\times l}$ is needed, the observers $\hat{c}_{ij}(j>l)$ are still indispensable. Without the connections between D_2 and R, i.e., without the third part of the right end of (2), the response network cannot realize partial topology identification, in general, since D_1 and R generally cannot reach the outer synchronization in this case. We have verified this assertion by simulation, but the simulation result is not shown here due to space constraint. The role of the connections between D_2 and R is to make D_1 and R symmetric in the structure, so that the outer synchronization between D_1 and R is possible.

In order to identify the partial topology that we need, the following assumptions and lemmas are necessary.

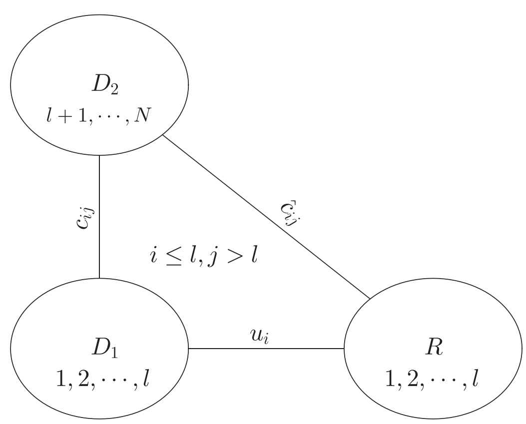

FIG. 1. The way how the response network connects with the drive network. D_1 denotes the subnetwork consisting of the first l nodes of the drive network, D_2 denotes the subnetwork consisting of the other N–l nodes, and R denotes the response network.

Assumption 1: (A1) Suppose that there exist nonnegative constants α and β satisfying

$$\|f_i(t,x_1) - f_i(t,x_2)\| \le \alpha \|x_1 - x_2\| \tag{3}$$

and

$$\|f_i(t_1,x) - f_i(t_2,x)\| \le \beta |t_1 - t_2| \tag{4}$$

for any $t, t_1, t_2 \in \mathbb{R}_+$ and any vectors $x, x_1, x_2 \in \mathbb{R}^n$, where $1 \le i \le N$.

Assumption 2: (A2) Suppose that $\{\Gamma x_i(t)\}_{i=1}^N$ are linearly independent on the orbit $\{x_i(t)\}_{i=1}^N$ of the outer synchronization manifold $\{x_i(t) = y_i(t)\}_{i=1}^l$.

Lemma 1: Let $g : \mathbb{R}_+ \to \mathbb{R}^m$ be differentiable, where m is a positive integer. If

$$\lim_{t\to+\infty} g(t) = 0,$$

and $\dot g$ is Lipschitz continuous, then $\lim_{t\to+\infty} \dot g(t) = 0$.

Proof: Since $\dot g$ is Lipschitz continuous, there exists such a positive constant L that

$$\|\dot g(t_1) - \dot g(t_2)\| < L|t_1 - t_2|, \quad \forall t_1,\, t_2 \in \mathbb{R}_+.$$

Denote $g = (g_1, g_2, \ldots g_m)^T$. Suppose $\lim_{t\to+\infty} \dot g_1(t) = 0$ does not hold. There exists a positive constant ε and a sequence $\{\omega_n\}$ in strictly increasing order satisfying that $\lim_{n\to+\infty} \omega_n = +\infty$ and $|\dot g_1(\omega_n)| > 2\sqrt{L\varepsilon}$. For $t \in \left(0, \sqrt{\frac{\varepsilon}{L}}\right)$, one has

$$|\dot g_1(\omega_n + t) - \dot g_1(\omega_n)| \le \|\dot g(\omega_n + t) - \dot g(\omega_n)\| < Lt \le \sqrt{L\varepsilon},$$

yielding that $|\dot g_1(\omega_n + t)| > \sqrt{L\varepsilon}$. It follows that

$$\left| g_1\left(\omega_n + \sqrt{\frac{\varepsilon}{L}}\right) - g_1(\omega_n)\right| = \sqrt{\frac{\varepsilon}{L}} \times |\dot g_1(\omega_n + \xi_n)| > \varepsilon, \tag{5}$$

where $\xi_n \in \left(0, \sqrt{\frac{\varepsilon}{L}}\right)$. Since $\lim_{t\to+\infty} g(t) = 0$, it is obvious that $\lim_{t\to+\infty} g_1(t) = 0$. There exists a $T \in \mathbb{R}_+$, satisfying that

$$|g_1(t_1) - g_1(t_2)| < \varepsilon, \quad \forall t_1, t_2 > T_1,$$

which contradicts with (5). Therefore, one deduces that $\lim_{t\to+\infty} \dot g_1(t) = 0$. Similarly, one has $\lim_{t\to+\infty} \dot g_k(t) = 0$ for $2 \le k \le m$, yielding that $\lim_{t\to+\infty} \dot g(t) = 0$. This completes the proof.

Lemma 2:[33] Assume that P is a diagonal matrix in which the first r diagonal elements are p and the others are 0, where $p > 0$ is a proper constant which is sufficiently large. Then, $G - P < 0$ is equivalent to $G(r) < 0$, where $G(r)$ denotes the minor matrix of matrix G by removing the first r rows and the first r columns of G.

III. MAIN RESULTS

Denoting $e_i(t) = y_i(t) - x_i(t), \tilde c_{ij} = \hat c_{ij} - c_{ij}$, the error system can be deduced from (1) and (2), described by

$$\dot e_i(t) = f_i(t, y_i(t)) - f_i(t, x_i(t)) + \sum_{j=1}^{l} c_{ij}\Gamma e_j(t) + \sum_{j=1}^{l} \tilde c_{ij}\Gamma y_j(t) + \sum_{j=l+1}^{N} \tilde c_{ij}\Gamma x_j(t) + u_i, \tag{6}$$

where $1 \le i \le l$, $1 \le j \le N$. Based on the assumptions (A1) and (A2), the following theorem can be deduced.

Theorem 1: Suppose (A1) and (A2) hold. Then, the partial topology C_{lxN} can be estimated by the estimation matrix $\hat C_{l\times N}$ via response network (2) with the following controllers and updating laws

$$\begin{aligned} u_i &= -d_i e_i(t), \\ \dot d_i &= k_i e_i^T(t) e_i(t), \\ \dot{\hat c}_{ij} &= -e_i^T(t)\Gamma y_j(t), \quad 1 \le j \le l, \\ \dot{\hat c}_{ij} &= -e_i^T(t)\Gamma x_j(t), \quad j > l, \end{aligned} \tag{7}$$

where $1 \le i \le l$ and k_i is any positive constant.

Proof: Since (A1) holds, one has

$$\|f_i(t, y_i(t)) - f_i(t, x_i(t))\| \le \alpha \|e_i(t)\|,$$

where $1 \le i \le l$. Choose a Lyapunov candidate as

$$V(t) = \frac{1}{2}\sum_{i=1}^{l} e_i^T(t) e_i(t) + \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{N} \tilde c_{ij}^2 + \frac{1}{2}\sum_{i=1}^{l} \frac{(d_i - d^*)^2}{k_i}, \tag{8}$$

where d^* is a positive constant to be determined. Then, one has

$$\begin{aligned} \dot V(t) &= \sum_{i=1}^{l} e_i^T(t)\dot e_i(t) + \sum_{i=1}^{l}\sum_{j=1}^{N} \tilde c_{ij}\dot{\hat c}_{ij} + \sum_{i=1}^{l} \frac{(d_i - d^*)\dot d_i}{k_i} \\ &\le \alpha \sum_{i=1}^{l} e_i^T(t) e_i(t) + \sum_{i=1}^{l}\sum_{j=1}^{l} c_{ij} e_i^T(t)\Gamma e_j(t) \\ &\quad + \sum_{i=1}^{l}\sum_{j=1}^{l} \tilde c_{ij} e_i^T(t)\Gamma y_j(t) + \sum_{i=1}^{l}\sum_{j=l+1}^{N} \tilde c_{ij} e_i^T(t)\Gamma x_j(t) \\ &\quad - \sum_{i=1}^{l} d_i e_i^T(t) e_i(t) + \sum_{i=1}^{l}\sum_{j=1}^{l} \tilde c_{ij}\left(-e_i^T(t)\Gamma y_j(t)\right) \\ &\quad + \sum_{i=1}^{l}\sum_{j=l+1}^{N} \tilde c_{ij}\left(-e_i^T(t)\Gamma x_j(t)\right) \\ &\quad + \sum_{i=1}^{l} (d_i - d^*) e_i^T(t) e_i(t) \\ &= (\alpha - d^*)\sum_{i=1}^{l} e_i^T(t) e_i(t) + \sum_{i=1}^{l}\sum_{j=1}^{l} c_{ij} e_i^T(t)\Gamma e_j(t). \end{aligned}$$

Denote $e(t) = (e_1^T(t), e_2^T(t), \ldots, e_l^T(t))^T \in \mathbb{R}^{nl}$ and $P = C_{l\times l} \otimes \Gamma$. Then, one gets

$$\begin{aligned}\dot{V}(t) &\le (\alpha - d^*)e^T(t)e(t) + e^T(t)Pe(t)\\ &= (\alpha - d^*)e^T(t)e(t) + e^T(t)\left(\frac{P+P^T}{2}\right)e(t)\\ &\le \left(\alpha - d^* + \frac{1}{2}\lambda_{max}(P+P^T)\right)e^T(t)e(t).\end{aligned}$$

Taking $d^* = \alpha + \frac{1}{2}\lambda_{max}(P+P^T) + 1$, one gets $\dot{V}(t) \le -e^T(t)e(t)$, from which it can be deduced that

$$\int_0^t e^T(s)e(s)\mathrm{d}s \le -\int_0^t \dot{V}(s)\mathrm{d}s = V(0) - V(t) \le V(0)$$

for any $t \ge 0$. Therefore, one has $e(t) \in L^2$. Since $\dot{V}(t) \le 0$, V is bounded. It follows from (8) that $e_i, \tilde{c}_{ij}$, and d_i are bounded, where $1 \le i \le l$ and $1 \le j \le N$. Due to the boundedness of network (1), x_j and $\dot{x}_j$ are bounded. Since $y_i = x_i + e_i$ is bounded, it follows from error system (6) that $\dot{e}_i$ is bounded. By using Barbalat lemma,[34] one obtains

$$\lim_{t\to+\infty} e_i(t) = 0. \tag{9}$$

Rewrite (6) as $\dot{e}_i(t) = \varphi_i(t) + \psi_i(t)$, where

$$\varphi_i(t) = f_i(t, y_i(t)) - f_i(t, x_i(t)) \tag{10}$$

and

$$\psi_i(t) = \sum_{j=1}^{l} c_{ij}\Gamma e_j(t) + \sum_{j=1}^{l} \tilde{c}_{ij}\Gamma y_j(t) + \sum_{j=l+1}^{N} \tilde{c}_{ij}\Gamma x_j(t) + u_i. \tag{11}$$

It is first to prove that ψ_i is Lipschitz continuous. From (11), one gets

$$\begin{aligned}\dot{\psi}_i(t) = &\sum_{j=1}^{l} c_{ij}\Gamma \dot{e}_j(t) + \sum_{j=1}^{l} \dot{\tilde{c}}_{ij}\Gamma y_j(t) + \sum_{j=1}^{l} \tilde{c}_{ij}\Gamma \dot{y}_j(t)\\ &+ \sum_{j=l+1}^{N} \dot{\tilde{c}}_{ij}\Gamma x_j(t) + \sum_{j=l+1}^{N} \tilde{c}_{ij}\Gamma \dot{x}_j(t)\\ &- \dot{d}_i e_i(t) - d_i \dot{e}_i(t).\end{aligned}$$

From (7) and $\dot{y}_i = \dot{x}_i + \dot{e}_i$, one deduces that $\dot{d}_i, \dot{\tilde{c}}_{ij}$ and $\dot{y}_i$ are bounded. It follows that $\dot{\psi}_i$ is bounded, yielding that ψ_i is Lipschitz continuous. Now, it is to prove that φ_i is Lipschitz continuous. From (10), one gets

$$\begin{aligned}|\varphi_i(t_1) - \varphi_i(t_2)| &= |f_i(t_1, y_i(t_1)) - f_i(t_1, x_i(t_1))\\ &\quad - f_i(t_2, y_i(t_2)) + f_i(t_2, x_i(t_2))|\\ &\le |f_i(t_1, x_i(t_1)) - f_i(t_1, x_i(t_2))|\\ &\quad + |f_i(t_1, x_i(t_2)) - f_i(t_2, x_i(t_2))|\\ &\quad + |f_i(t_1, y_i(t_1)) - f_i(t_1, y_i(t_2))|\\ &\quad + |f_i(t_1, y_i(t_2)) - f_i(t_2, y_i(t_2))|\\ &\le \alpha|x_i(t_i) - x_i(t_2)| + \alpha|y_i(t_1) - y_i(t_2)|\\ &\quad + 2\beta|t_1 - t_2|\\ &= \big(\alpha|\dot{x}_i(\xi_1)| + \alpha|\dot{y}_i(\xi_2)| + 2\beta\big)|t_1 - t_2|,\end{aligned}$$

where $1 \le i \le l$, $t_1 < t_2$ and $\xi_1, \xi_2 \in (t_1, t_2)$. Due to the boundedness of $\dot{x}_i$ and $\dot{y}_i$, one gets that φ_i is Lipschitz continuous. Therefore, $\dot{e}_i$ is Lipschitz continuous. According to Lemma 1, one has

$$\lim_{t\to+\infty} \dot{e}_i(t) = 0. \tag{12}$$

It then follows from (6) that

$$\lim_{t\to+\infty}\left(\sum_{j=1}^{l} \tilde{c}_{ij}\Gamma y_j(t) + \sum_{j=l+1}^{N} \tilde{c}_{ij}\Gamma x_j(t)\right) = 0. \tag{13}$$

Furthermore, one has

$$\lim_{t\to+\infty}\sum_{j=1}^{N} \tilde{c}_{ij}\Gamma x_j(t) = 0, \quad 1 \le i \le l. \tag{14}$$

From (A2), one has $\lim_{t\to+\infty}\tilde{c}_{ij} = 0$, $1 \le i \le l$, $1 \le j \le N$, which indicates that the partial topology $C_{l\times N}$ can be identified by the estimation matrix $\hat{C}_{l\times N}$. This completes the proof.

Remark 2: The proof of (12) is neglected in some studies.[22,27] Besides (9), the Lipschitz continuity of $\dot{e}_i$ is also a critical condition guaranteeing (12). Consider a function

$$g(t) = \frac{1}{t+1}\sin e^t, \quad t \in \mathbb{R}_+.$$

It is obvious that $\lim_{t\to+\infty}g(t) = 0$, but one does not have $\lim_{t\to+\infty}\dot{g}(t) = 0$.

Remark 3: When $l = N$, the topology of the whole network can be identified. Therefore, our method is able to deal with not only partial topology but also the whole topology of complex networks.

Remark 4: Compared with traditional methods,[18,19,23] our method is of great advantage when we are interested in only a small part of the whole topology. For example, we would like to know the couplings between the first three nodes and their neighbors in a complex network composed of $N = 1000$ nodes. By using our method, we only need to construct $l = 3$ nodes and lN observers to estimate the unknown couplings. While by using traditional methods, we need to construct N nodes and N^2 observers, which is too consuming to be practical. From another perspective, the dimension of the response network with controllers and updating laws (7) is $l(n+1+N)$, while that in Refs. 18, 19, and 23 is $N(n+1+N)$. Therefore, our method is more cost-effective and provides an approach for dealing with large-scale networks.

Let $l = 1$, then we have the following corollary.

Corollary 1: Suppose (A1) and (A2) hold. Then, the connections starting from node 1 can be identified via a response network

$$\dot{y}_1(t) = f_1(t, y_1(t)) + \hat{c}_{11}\Gamma y_1(t) + \sum_{j=2}^{N} \hat{c}_{1j}\Gamma x_j(t) + u_1, \tag{15}$$

with the following controllers and updating laws

$$
\begin{aligned}
&u_1 = -d_1 e_1(t),\\
&\dot{d}_1 = k_1 e_1^T(t) e_1(t),\\
&\dot{\hat{c}}_{11} = -e_1^T(t)\Gamma y_1(t),\\
&\dot{\hat{c}}_{1j} = -e_1^T(t)\Gamma x_j(t), j > 1,
\end{aligned}
$$

where k_1 is any positive constant.

Remark 5: From corollary 1, all the connections starting from node 1 can be identified via a response network consisting of just one node (though there is just one node, we still refer to it as network). For a drive network of identical nodes, we can renumber any node as node 1 through rearranging the ordinal of the nodes, thus the connections starting from this node can be identified via response network (15). Therefore, constructing only one node and N observers, $\hat{C}_{l\times N}$ can be identified by using (15) l times. Particularly, the topology of the whole network can be identified by using (15) N times.

Sometimes, only some couplings of the drive network are unknown, while the others have already been known. In this case, the constructed response network is simpler and more cost-effective.

Corollary 2: Suppose (A1) and (A2) hold. If only partial couplings $c_{ij}(1 \le i \le l,\ 1 \le j \le l_1,\ l \le l_1)$ are unknown, then $C_{l\times l_1} = (c_{ij})_{l\times l_1}$ can be estimated by the estimation matrix $\hat{C}_{l\times l_1} = (\hat{c}_{ij})_{l\times l_1}$ via response network (2) with the following controllers and updating laws

$$
\begin{aligned}
&u_i = -d_i e_i(t),\\
&\dot{d}_i = k_i e_i^T(t) e_i(t),\\
&\dot{\hat{c}}_{ij} = -e_i^T(t)\Gamma y_j(t), \quad 1 \le j \le l,\\
&\dot{\hat{c}}_{ij} = -e_i^T(t)\Gamma x_j(t), \quad l < j \le l_1,\\
&\hat{c}_{ij} = c_{ij}, \quad j > l_1,
\end{aligned}
\tag{16}
$$

where $1 \le i \le l$, and k_i is any positive constant.

Proof: Choose a Lyapunov candidate as

$$
V(t) = \frac{1}{2}\sum_{i=1}^{l} e_i^T(t)e_i(t) + \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l_1}\tilde{c}_{ij}^2 + \frac{1}{2}\sum_{i=1}^{l}\frac{(d_i - d^*)^2}{k_i},
$$

where d* is a positive constant to be determined. The rest of the proof is very similar to that of Theorem 1 and thus omitted here.

Remark 6: Reference 26 also studied the case that only a part of the couplings are unknown. The dimension of the response network therein is $nN + l + l \times l_1$, while the dimension of the response network with controllers and updating laws (16) is $nl + l + l \times l_1$. Therefore, the response network with controllers and updating laws (16) is simpler and more cost-effective than that in Ref. 26.

Suppose $C_{l\times l}$ has already been known, and the goal is to identify the couplings between the first l nodes and the other N–l nodes. In this case, the number of controllers u_i can be less than l. Denote $\gamma = \|\Gamma\|, \kappa = \frac{1}{2}\lambda_{\min}(\Gamma + \Gamma^T)$, and $\bar{C} = \frac{1}{2}(\check{C} + \check{C}^T)$, where $\check{C}$ is a modified matrix of $C_{l\times l}$ by replacing c_{ii} with $\frac{\kappa}{\gamma}c_{ii}$.

Corollary 3: Suppose (A1) and (A2) hold, and $C_{l\times l}$ is known. If there exists a positive integer r satisfying $1 \le r < l$ and $\lambda_{\max}(\bar{C}(r)) < -\frac{\alpha}{\gamma}$, then the couplings $c_{ij}(1 \le i \le l,\ l < j \le N)$ can be estimated by $\hat{c}_{ij}$ via response network (2) with the following controllers and updating laws

$$
\begin{aligned}
&u_i = -d_i e_i(t), \quad 1 \le i \le r,\\
&u_i = 0, \quad r < i \le l,\\
&\dot{d}_i = k_i e_i^T(t) e_i(t), \quad 1 \le i \le r,\\
&\hat{c}_{ij} = c_{ij}, \quad 1 \le i, j \le l,\\
&\dot{\hat{c}}_{ij} = -e_i^T(t)\Gamma x_j(t), \quad 1 \le i \le l, \quad j > l \quad \text{or} \quad l < j
\end{aligned}
\tag{17}
$$

where k_i is any positive constant.

Proof: Choose a Lyapunov candidate as

$$
V(t) = \frac{1}{2}\sum_{i=1}^{l} e_i^T(t)e_i(t) + \frac{1}{2}\sum_{i=1}^{l}\sum_{j=l+1}^{N}\tilde{c}_{ij}^2 + \frac{1}{2}\sum_{i=1}^{r}\frac{(d_i - d^*)^2}{k_i},
$$

where d^* is a positive constant to be determined. Then, one has

$$
\begin{aligned}
\dot{V}(t) &\le \alpha\sum_{i=1}^{l} e_i^T(t)e_i(t) + \sum_{i=1}^{l}\sum_{j=1}^{l} c_{ij} e_i^T(t)\Gamma e_j(t)\\
&\quad - d^*\sum_{i=1}^{r} e_i^T(t)e_i(t)\\
&\le \alpha\sum_{i=1}^{l} e_i^T(t)e_i(t) + \gamma\sum_{i=1}^{l}\sum_{j=1, j\ne i}^{l} c_{ij}\|e_i(t)\|\|e_j(t)\|\\
&\quad + \kappa\sum_{i=1}^{l} c_{ii} e_i^T(t)e_i(t) - d^*\sum_{i=1}^{r} e_i^T(t)e_i(t)\\
&= \mu^T(t)\big(\alpha I_l + \gamma\bar{C} - D\big)\mu(t),
\end{aligned}
$$

where $\mu(t) = (\|e_1(t)\|, \ldots, \|e_l(t)\|)^T$ and $D \in \mathbb{R}^{l\times l}$ is a diagonal matrix in which the first r diagonal elements are d^* and the others are 0. Since $\lambda_{\max}(\bar{C}(r)) < -\frac{\alpha}{\gamma}$, one has $\alpha I_{l-r} + \gamma\bar{C}(r) < 0$. According to Lemma 2, one can select a sufficiently large d^* satisfying that $\alpha I_l + \gamma\bar{C} - D < 0$. The rest of the proof is very similar to that of Theorem 1 and thus omitted here.

Remark 7: (A2) is a critical condition for guaranteeing the identification success. However, it is very difficult to verify (A2) in spite of its importance. To the best of our knowledge, there has not yet been an effective method for verifying (A2). We give here two necessary but not sufficient conditions for (A2). Suppose the identified topology is $C^* = \lim_{t\to+\infty}\hat{C}_{l\times N}(t) \in \mathbb{R}^{l\times N}$. If (A2) holds, C^* should satisfy the following two conditions:

(C1) $c_{ij}^* \ge 0, 1 \le i \le l, 1 \le j \le N, i \ne j;$
(C2) for the undirected drive network, $c_{ij}^* = c_{ji}^*, 1 \le i, j \le l.$

Finding a simple and sufficient condition for (A2) is still a challenging problem.

Remark 8: The condition $\sum_{j=1}^{N} c_{ij}^* = 0 (1 \le i \le l)$ is also necessary for (A2), but it is of little use. When the identification fails, we know that (A2) does not hold. In this case, network (1) usually reaches complete synchronization, that is,

$$x_1(t) = x_2(t) = \cdots = x_N(t).$$

Then, (14) turns into

$$\lim_{t\to+\infty} (\tilde{c}_{i1} + \tilde{c}_{i2} + \cdots + \tilde{c}_{iN}) = 0,\ 1 \le i \le 1,$$

yielding that $\sum_{j=1}^{N} c_{ij}^* = 0$. Therefore, $\sum_{j=1}^{N} c_{ij}^* = 0 (1 \le i \le l)$ usually holds, no matter whether (A2) holds.

IV. NUMERICAL SIMULATION

The Lü system is a well-known benchmark chaotic system[35] described by

$$\dot{x} = g(x) = Bx + W(x), \tag{18}$$

where

$$B = \begin{pmatrix} -a & a & 0 \\ 0 & c & 0 \\ 0 & 0 & -b \end{pmatrix},\quad W(x) = \begin{pmatrix} 0 \\ -x_1x_3 \\ x_1x_2 \end{pmatrix},$$

with $a=36$, $b=3$, and $c=20$. It can be simply verified that the Lü system satisfies (A1) due to its boundedness.[18,35]

Next, we present three networks in order to verify the effectiveness of the proposed method. In all the three networks, we are always interested in the couplings among three nodes, say 1, 2, and 3 (if they are not numbered 1, 2, and 3, we can rearrange the ordinal of the nodes so that they are numbered 1, 2, and 3). Then, we have $l=3$. Moreover, we always use response network (2) with controllers and updating laws (7) to identify the partial topology in which we are interested.

A. A 20-node small-world network of identical nodes

Consider a network consisting of $N=20$ Lü systems, described by

$$\dot{x}_i(t) = g(x_i(t)) + \sum_{j=1}^{20} c_{ij}\Gamma x_j(t), \tag{19}$$

where $1 \le i \le 20$, $\Gamma = \begin{pmatrix} 1 & 0 & 0.5 \\ 0 & 1 & 0.3 \\ 0.4 & 0 & 1 \end{pmatrix}$, and $C = \varepsilon C_{\text{adj}}$ with the coupling strength $\varepsilon > 0$. The non-diagonal elements of C_{adj} are depicted in Fig. 2, while the diagonal elements should be calculated according to the zero row sum. The Watts-Strogatz (W-S) algorithm[4] is used here to generate a small-world network. First, start with a ring of 20 nodes, each connected to its $K=4$ nearest neighbors. The coupling weight is an integer between 1 and 4, which is generated randomly and illustrated by the corresponding color, as shown in Fig. 2. Second, rewire each edge with probability $p_r=0.2$ in a certain way.[4]

From Fig. 3, we can see that the identification of the partial topology succeeds when $\varepsilon=0.1$ but fails when $\varepsilon=1$. This is because Γ is anti-stable. When Γ is anti-stable, network (19) is of type I, i.e., its synchronization region is unbounded.[36] In this case, large coupling strength favors synchronization within the drive network and thus hinders the identification. Without comparing with the original topology, we can still know that the identification fails when $\varepsilon=1$, since the identified partial topology in Fig. 3(b) satisfies neither (C1) nor (C2) in Remark 7.

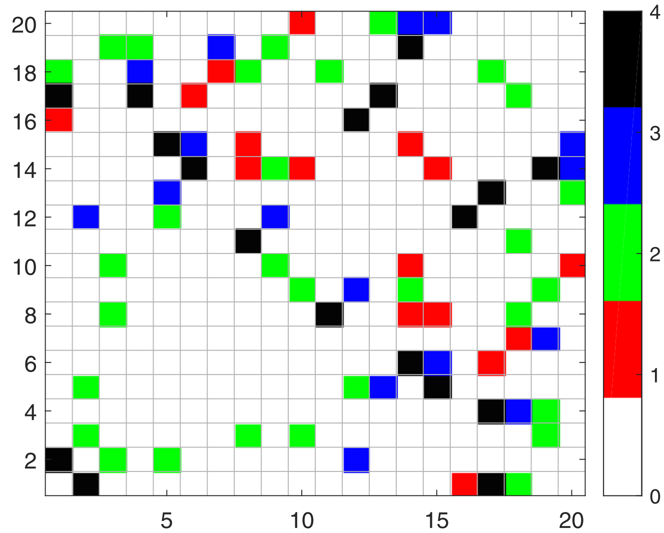

FIG. 2. Colormap of a 20-node W-S small-world network generated with rewiring probability $p_r=0.2$.

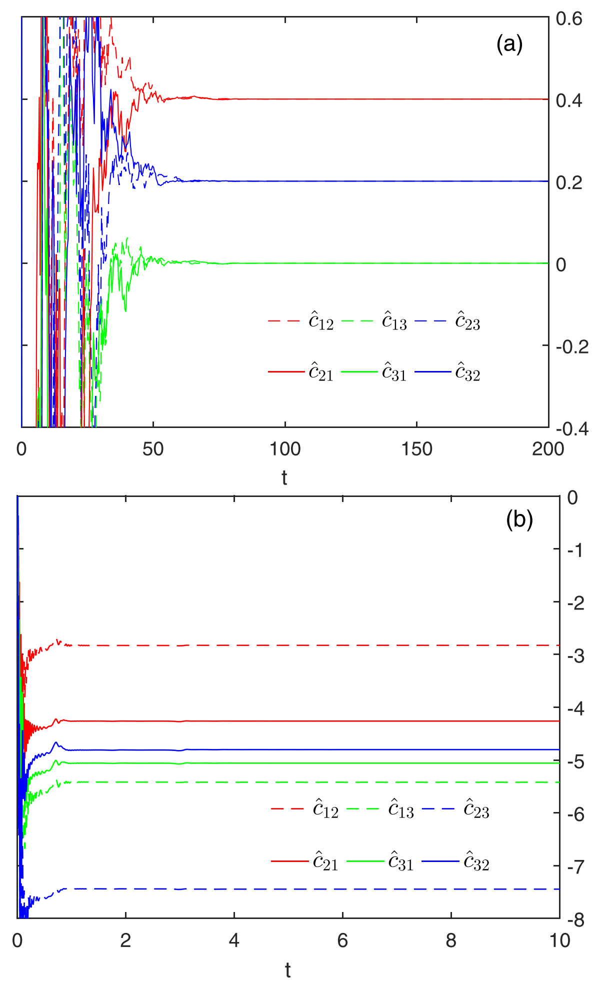

FIG. 3. Partial topology identification of network (19) with (a) $\varepsilon=0.1$ or (b) $\varepsilon=1$: evolution of the coupling set $\{\hat{c}_{12}, \hat{c}_{13}, \hat{c}_{21}, \hat{c}_{23}, \hat{c}_{31}, \hat{c}_{32}\}$.

Some initial values and parameters are given as follows, $d_i(0)=1$, $k_i=1$, $x_j(0)=3j+(1,\ 2,\ 3)^T$, $y_i(0)=-2.6+0.6i+0.2(1,\ 2,\ 3)^T$, and $\hat{c}_{ij}(0)=1$, where $1\le i\le 3$ and $1\le j\le 20$.

B. A 20-node small-world network of nonidentical nodes

As Ref. 37 stated, synchronization within the drive network is an obstacle to the identification of network topology, since (A2) does not hold when the drive network reaches synchronization. Even generalized synchronization may hinder the identification. Many complex networks with nonidentical nodes naturally satisfy (A2).[22] The big node difference can shorten the time that the identification needs and improve the identification efficiency.[38] Therefore, a simple method for avoiding identification failure is to make the nodes nonidentical. Add artificial perturbation to the Lü system, say, $v_i=\rho\cos(it)(1,1,1)^T$, where $i\ge 1$ and ρ is any positive constant. Then, we get a class of different systems described by

$$f_i(t,x(t)) = g(x(t)) + \rho\cos(it)(1,1,1)^T, \quad i\ge 1. \tag{20}$$

It is clear that the difference between f_i and f_i is bigger with a larger value of ρ, where $i\neq j$ and $i,j\ge 1$.

Consider a network consisting of $N=20$ nonidentical nodes, described by

$$\dot{x}_i(t) = f_i(t,x_i(t)) + \sum_{j=1}^{20} c_{ij}\Gamma x_j(t), \tag{21}$$

where $1\le i\le 20$, and Γ, $C=\varepsilon C_{\mathrm{adj}}$ are the same as those in the last example. Since

$$\|f_i(t,x)-f_i(t,y)\| = \|g(x)-g(y)\|,$$

and

$$\begin{aligned}\|f_i(t_1,x)-f_i(t_2,x)\| &= \sqrt{3}\rho|\cos(it_1)-\cos(it_2)|\\ &\le 20\sqrt{3}\rho|t_1-t_2|,\end{aligned} \tag{22}$$

(A1) holds.

The identification fails when $\rho=0$ according to the last example (the initial values and parameters are the same as those in the last example). However, from Fig. 4, we can see that the identification succeeds when $\rho=10$, 100, and 1000, indicating that nonidentical nodes favors the identification. Specifically, we can see from Fig. 4 that the time needed for identification decreases significantly with the increase of ρ, which is consistent with the result in Ref. 38. To make the node difference bigger, we can simply increase the value of ρ, i.e., add stronger perturbation. However, too strong perturbation may not be practical and may destroy the original network, so we should choose the perturbation of appropriate strength.

C. A 1000-node ring network of nonidentical nodes

Consider a ring network consisting of $N=1000$ nonidentical nodes, described by

$$\dot{x}_i(t) = f_i(t,x_i(t)) + \varepsilon\Gamma(x_{i-1}(t)-2x_i(t)+x_{i+1}(t)), \tag{23}$$

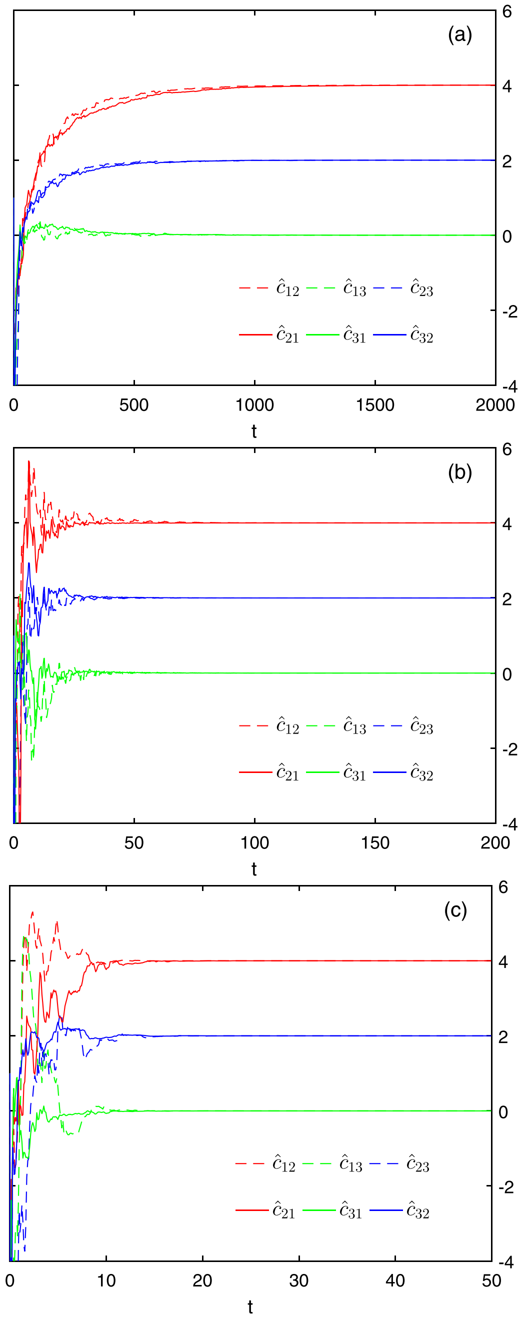

FIG. 4. Partial topology identification of network (21) with (a) $\rho=10$, (b) $\rho=100$ or (c) $\rho=1000$: evolution of the coupling set $\{\hat{c}_{12},\hat{c}_{13},\hat{c}_{21},\hat{c}_{23},\hat{c}_{31},\hat{c}_{32}\}$.

where $1\le i\le 1000$, $\Gamma=I_3$, $x_0\equiv x_{1000}$, $x_{1001}\equiv x_1$ and $\varepsilon>0$ is the coupling strength. According to the last example, (A1) holds. Figure 5 shows that the partial topology is correctly identified.

Some initial values and parameters are given as follows: $\varepsilon=0.2$, $\rho=20$, $d_i(0)=1$, $k_i=1$, $x_j(0)=1+0.001j+0.002(1,\ 2,\ 3)^T$, $y_i(0)=1+i+(1,\ 2,\ 3)^T$, and $\hat{c}_{ij}(0)=1$, where $1\le i\le 3, 1\le j\le 1000$.

V. CONCLUSION

In this paper, we have proposed a novel and cost-effective method for identifying the partial topology of

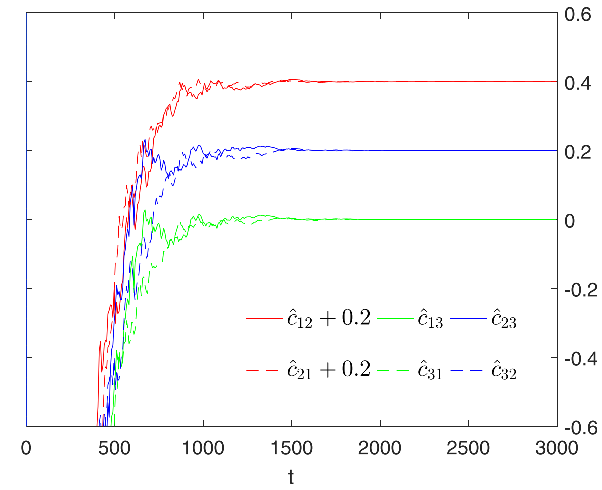

FIG. 5. Partial topology identification of network (23) evolution of the coupling set $\{\hat{c}_{12}, \hat{c}_{13}, \hat{c}_{21}, \hat{c}_{23}, \hat{c}_{31}, \hat{c}_{32}\}$.

complex dynamical networks via a pinning mechanism. Compared with most of the existing methods, our method is able to deal with large-scale networks and is more cost-effective when we are only interested in only a small part of the whole topology, because the number of nodes and observers in the response network are greatly reduced. Several numerical examples are provided to verify the effectiveness of the proposed method. For networks of identical nodes, the identification may fail. To avoid identification failure, we present a simple method for making the nodes nonidentical by adding artificial perturbation, such as $v_i = \rho\cos(it)(1,1,1)^T$. We introduce two simple and necessary conditions for (A2), while finding a simple and sufficient condition for (A2) remains an important and challenging problem. In 2017, Ref. 39 proposed a compressive-sensing-based method for identifying one layer (it is, in fact, partial topology) of multilayer networks. However, corresponding methods based on synchronization are rare. We expect to extend our work to the context of multilayer networks in further study.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China under Grant Nos. 61374173, 61773294, and 61773175.

[1]R. Albert and A. L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[2]S. H. Strogatz, Nature **410**, 268 (2001).
[3]M. E. J. Newman, SIAM Rev. **45**, 167–256 (2003).
[4]D. J. Watts and S. H. Strogatz, Nature **393**, 440 (1998).
[5]A. L. Barabási and R. Albert, Science **286**, 509–512 (1999).
[6]Z. Tang, H. P. Ju, and J. Feng, Commun. Nonlinear Sci. Numer. Simul. **57**, 422–438 (2018).
[7]R. Albert, H. Jeong, and A. L. Barabási, Nature **406**, 378–382 (2000).
[8]M. E. J. Newman and D. J. Watts, Phys. Rev. E **60**, 7332 (1999).
[9]C. Yin, S. Zhu, H. Chen, B. Zhang, and B. David, Int. J. Distrib. Sens. Networks **2015**, 849140.
[10]Z. Tang, J. H. Park, and W. X. Zheng, Int. J. Robust Nonlinear Control **28**(3), 1001–1015 (2018).
[11]J. Zhou, J. Lu, and J. Lii, Automatica **44**, 996 (2008).
[12]L. M. Pecora and T. L. Carroll, Int. J. Bifurcation Chaos **9**, 2315 (1999).
[13]J. Zhou, J. Lu, and J. Lü, IEEE Trans. Autom. Control **51**, 652 (2006).
[14]Y. Wu and C. Li, Commun. Nonlinear Sci. Numer. Simul. **17**, 349 (2012).
[15]Z. Tang, H. P. Ju, and T. H. Lee, Commun. Nonlinear Sci. Numer. Simul. **39**, 7 (2016).
[16]H. Shi, Y. Sun, L. Miao, and Z. Duan, Nonlinear Dyn. **85**, 2437 (2016).
[17]J. Zhou, J. Chen, J. Lu, and J. Lü, IEEE Trans. Autom. Control **62**, 3468 (2017).
[18]J. Zhou and J. Lu, Physica A **386**, 481 (2007).
[19]X. Wu, Physica A **387**, 997 (2008).
[20]Y. Che, R. X. Li, C. X. Han, J. Wang, S. G. Cui, B. Deng, and X. Wei, Eur. Phys. J. B **85**, 265 (2012).
[21]Q. Zhang, J. Luo, and L. Wan, Nonlinear Dyn. **71**, 353 (2013).
[22]H. Liu, J. Lu, J. Lü, and D. J. Hill, Automatica **45**, 1799 (2009).
[23]X. Wu, X. Zhao, J. Lü, L. Tang, and J. Lu, IEEE Trans. Control Network Syst. **3**, 379 (2016).
[24]H. Zhao, L. Li, H. Peng, J. Xiao, Y. Yang, and M. Zheng, Nonlinear Dyn. **83**, 1437–1451 (2016).
[25]S. Wang, M. Jiang, J. Mei, and J. Han, in *International Conference on Intelligent Human Machine System and Cybernetics* (2013) pp. 513–516.
[26]J. Zhou, W. Yu, X. Li, M. Small, and J. Lu, IEEE Trans. Neural Network **20**, 1679 (2009).
[27]S. Zhang, X. Wu, J. Lu, H. Feng, and J. Lü, IEEE Trans. Circuits Syst. I, Regul. Pap. **61**, 3216 (2014).
[28]Z. Tang, H. P. Ju, and T. H. Lee, Nonlinear Dyn. **85**, 2171 (2016).
[29]Y. Xu, W. Zhou, J. Zhang, W. Sun, and D. Tong, Nonlinear Dyn. **88**, 2969 (2017).
[30]J. Zhao, Q. Li, J. Lu, and Z. P. Jiang, Chaos **20**, 023119 (2010).
[31]K. I. Goh, M. E. Cusick, D. Valle, B. Childs, M. Vidal, and A. L. Barabsi, Proc. Natl. Acad. Sci. U. S. A. **104**, 8685 (2007).
[32]H. Paik, H. S. Heo, H. Ban, and S. B. Cho, J. Transl. Med. **12**(1), 99 (2014).
[33]J. Zhou, X. Wu, W. Yu, M. Small, and J. A. Lu, Chaos **18**, 377 (2008).
[34]M. Sun, IEEE Trans. Autom. Control **54**, 2222 (2009).
[35]J. Lü and G. Chen, Int. J. Bifurcation Chaos **12**, 659 (2002).
[36]C. Liu, Z. Duan, G. Chen, and L. Huang, Physica A **386**, 531 (2007).
[37]L. Chen, J. Lu, and K. T. Chi, IEEE Trans. Circuits Syst. II **56**, 310 (2009).
[38]L. Tang, J. Lu, X. Wu, and J. Lü, Nonlinear Dyn. **73**, 1081 (2013).
[39]G. Mei, X. Wu, Y. Wang, M. Hu, J. A. Lu, and G. Chen, IEEE Trans. Cybern. **48**(2), 754–764 (2018).