Synchronization: An Obstacle to Identification of Network Topology

Liang Chen, Jun-an Lu, and Chi K. Tse, Fellow, IEEE

Abstract—Identification of network topology is an important issue for physics, biology, engineering, and other science disciplines. Under the application of an adaptive-feedback control algorithm, we explore a process through which the topology of an unknown dynamical network can be identified. Analytical results show that the key to guaranteeing successful topology identification is the existence of the condition that all coupling terms of the unknown network are linearly independent of the synchronization manifold between the drive (unknown) network and a response network. When complete synchronization is achieved in the unknown network, the linearly independent condition is no longer satisfied, making its topology unidentifiable. We also find that partial synchronization in the unknown network implies a part of topology being unidentifiable. The results can be extended to projection synchronization and some generalized synchronization. Furthermore, we show that when the network dynamical equation is stable, its topology can be identified subject to some limitations. Finally, a method of avoiding synchronization for a network is presented and verified by numerical simulations.

Index Terms—Adaptive-feedback control, network, synchronization, topology identification.

I. INTRODUCTION

TODAY, we live in a world full of complex networks. Examples of networks that are making crucial contributions to our lives include transportation, phone call network, the Internet, and the World Wide Web [1], [2]. Significant progress has been made in studying complex networks since the discovery of their small-world [3] and scale-free [4] characteristics. So far, much of the research on complex networks has focused on modeling, dynamical analysis (synchronization), control, and structure optimization [5]–[10]. Recently, an important topic has emerged, which can be described as a problem of identifying the topology of an unknown network. Indeed, among the many practical problems associated with complex networks, identifying network topologies is the most fundamental. For instance, if a major malfunction occurs in a communication network, power network, or the Internet, it is very important to quickly detect the location of the faulty line. There are also other works that involve network identification, such as understanding the protein–DNA interactions in cell processes [11]. Therefore, identifying the topologies of networks is crucial to solving many problems associated with real networks [12].

The scheme used for identifying the topology of a network in this brief is the well-known adaptive-feedback control method, which has been developed to identify the model parameters of a given dynamical system (a particular chaotic system) [13]–[16]. In [17], results from numerical considerations have shown that the parameters of a stable system cannot be estimated by using the adaptive method. In [18], the parameter estimation of the model investigated in [13] was revisited, and it has been shown that parameter estimation can only be done for periodic or chaotic systems. However, recent research has pointed out that the earlier results on parameter identification of a single dynamical system [13]–[18] were either incomplete or incorrect, and that the key to the successful estimation of parameters is a linear independence condition [19], [20]. In [19], the parameter identification problem of a single dynamical system was considered, whereas in this brief, we focus on the network topology identification. In particular, we study how the synchronization in a network would affect the process of identifying its topology. This important issue has not been discussed in [19] or other prior works.

Naturally, the adaptive-feedback control method can be extended to identifying the topology of a network [12], [21], [22]. We will show in this brief that the topology identification process based on [12], [21], and [22] would fail if the network is in a synchronous situation. Using an adaptive-feedback control algorithm, we explore a process whereby the topology of an unknown dynamical network can be identified. We analytically show the impact of synchronization on the identification process.

II. NETWORK IDENTIFICATION

To demonstrate the identification of a network, we consider a dynamical network consisting of $N$ coupled oscillators, with each node being an $n$-dimensional dynamical system described by

\[ \dot{x}_i = F_i(x_i) + \sum_{j=1}^{N} c_{ij} H(x_j), \quad i = 1, 2, \ldots, N \]  

(1)
where \( x_i = (x_{i1}, x_{i2}, \ldots, x_{in})^T \) is the state vector of the \( i \)th oscillator, the function \( F_i \) is the corresponding nonlinear vector field, and \( H(\cdot) \) is the linear or nonlinear output function of the individual oscillators. In addition, the coupling matrix \( C = (c_{ij})_{N \times N} \) describes the coupling topology of the network, in which \( c_{ij} \neq 0 \) if there is a coupling from oscillator \( i \) to \( j \) (\( i \neq j \)); otherwise, \( c_{ij} = 0 \). Here, we do not assume that \( C \) is symmetric, irreducible, or diffusive.

The topology information of a network can entirely be represented in its coupling matrix \( C \). Accordingly, the identification of the topology of a network, as given in (1), becomes a problem of estimating the elements of matrix \( C \). To formulate our solution, a useful introduction is assumed as follows. We assume that the vector functions \( F_i \) and \( H \) satisfy the Lipschitz condition, i.e., there exist positive constants \( L_H \) and \( L_i \) such that

\[
|F_i(x) - F_i(y)| \leq L_i|x - y|, \quad |H(x) - H(y)| \leq L_H|x - y|
\]

where \( \| \cdot \| \) denotes the Euclidean vector norm.

To identify the elements of matrix \( C \), one should construct a drive-response system. We take the network in (1) as the drive network, and the response network with adaptive-feedback law can be designed as

\[
\begin{align*}
\dot{y}_i &= F_i(y_i) + \sum_{j=1}^{N} d_{ij} H(y_j) + u_i \quad \text{(2a)} \\
u_i &= -k_i e_i \quad \text{(2b)} \\
k_i &= r_i \|e_i\|^2 \quad \text{(3)}
\end{align*}
\]

where \( e_i = y_i - x_i \) denotes the synchronous errors, \( u_i \) is the controller for oscillator \( i \), \( k_i \) and \( d_{ij} \) are the adaptive parameters of the response system [see (2)] in the adaptive-feedback algorithm [see (3)], and \( r_i \) is an arbitrary positive constant for \( i, j = 1, \ldots, N \).

In the following, we show how the unknown \( c_{ij} \) may dynamically be estimated from \( d_{ij} \) in the response system [see (2)]. The error system between system (1) and system (2) can be written as follows:

\[
\dot{e}_i = F_i(y_i) - F_i(x_i) + \sum_{j=1}^{N} (d_{ij} H(y_i) - c_{ij} H(x_i)) - k_i e_i.
\]

By choosing the Lyapunov candidate as

\[
V = \frac{1}{2} \sum_{i=1}^{N} e_i^T e_i + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (d_{ij} - c_{ij})^2 + \frac{1}{2} \sum_{i=1}^{N} \frac{1}{r_i} (k_i - \rho)^2
\]

where \( \rho \) is a sufficiently large positive constant to be determined, we get the differential coefficient of \( V \) along systems (3) and (4) as

\[
\begin{align*}
\dot{V} &= \sum_{i=1}^{N} e_i^T [F_i(y_i) - F_i(x_i) - k_i e_i] \\
&+ \sum_{i=1}^{N} \sum_{j=1}^{N} e_i^T [d_{ij} H(y_j) - c_{ij} H(x_j)] \\
&- \sum_{i=1}^{N} \sum_{j=1}^{N} (d_{ij} - c_{ij}) e_i^T H(y_j) + \sum_{i=1}^{N} (k_i - \rho) \|e_i\|^2 \\
&\leq \sum_{i=1}^{N} (L_i - \rho) \|e_i\|^2 + L_H \sum_{i=1}^{N} \sum_{j=1}^{N} \|c_{ij}\| \cdot \|e_i\| \cdot \|e_j\| \\
&= E^T Q E
\end{align*}
\]

where

\[
\begin{align*}
E &= ([e_1, \|e_2\|, \ldots, \|e_N\|])^T \in \mathbb{R}^N \\
Q &= (P + P^T)/2 \\
P &= \text{diag}(L_1 - \rho, L_2 - \rho, \ldots, L_N - \rho) + L_H \tilde{C}
\end{align*}
\]

in which \( \tilde{C} = (|c_{ij}|)_{N \times N} \). It is obvious that there exists a sufficiently large positive constant \( \rho \) such that \( Q \) is negative definite, namely, \( \dot{V} \leq 0 \). Let \( M \) be the set of all points where \( \dot{V} = 0 \), which is easily found from \( M = \{ (E, D, k) \| V = 0 \} = \{ (E, D, k) \| e_i = 0, i = 1, 2, \ldots, N \} \), where \( D = (d_{ij})_{N \times N} \), and \( k = (k_1, k_2, \ldots, k_N)^T \). From \( e_1 = 0 \Leftrightarrow x_1 = y_1 \), which yields \( k_1 = 0 \), \( d_{ij} = 0 \), and \( \sum_{j=1}^{N} (d_{ij} - c_{ij}) H(x_j) = 0 \), the largest invariant set of \( M \) can be written as

\[
\bar{M} = \left\{ (E, D, k) \| e_i = 0, k_i = 0, d_{ij} = 0, \sum_{j=1}^{N} (d_{ij} - c_{ij}) H(x_j) = 0, i, j = 1, 2, \ldots, N \right\}.
\]

According to the invariant principle of differential equations [23], starting with arbitrary initial values, the trajectory asymptotically converges to the largest invariant \( \bar{M} \), which implies that \( e_i \to 0 \), \( d_{ij} \to d_{ij}^* \), and \( k_i \to k_i^* \), where \( d_{ij}^* \) and \( k_i^* \) are constants depending on the initial values, for \( i, j = 1, 2, \ldots, N \), as \( t \to \infty \).

In [12], [21], and [22], it has been argued that the identification process completes at this point, and the topology can successfully be estimated. However, the above condition does not guarantee that \( d_{ij} \to c_{ij} \) as \( t \to \infty \). This is because when we obtain \( e_i \to 0 \) from the stable error system [see (4)], there may exist many nonzero parameters \( p_j \neq 0, j = 1, 2, \ldots, N \), such that \( \sum_{j=1}^{N} p_j H(x_j) = 0 \), and \( (d_{ij} - c_{ij}) = p_j \). Thus, further conditions need to be imposed for identification purposes. Specifically, a linear independence condition is proposed as follows.

The functions \( H(x_j) \) are linearly independent if and only if there do not exist nonzero constants \( \beta_j \) \( (j = 1, 2, \ldots, N) \) such that \( \beta_1 H(x_1) + \beta_2 H(x_2) + \cdots + \beta_N H(x_N) = 0 \).
When all functions \( H(x_j) \) satisfy the above \textit{linear independence condition} on the synchronization manifold \( \{e_i = 0\} \), then \( d_{ij} = c_{ij} \), i.e., the topology of the network of (1) is successfully identified.

Based on the \textit{linear independence condition}, we propose another useful and simple condition for accomplishing identification, as follows. Denoting \( H(x_j) \) as \( (h_1(x_j), h_2(x_j), \ldots, h_n(x_j))^T \), where \( h_i \) is a subfunction of \( H(x_j) \), if there is a positive integer \( \alpha \in \{1, 2, \ldots, n\} \) such that the subfunctions \( h_1(x_j), h_2(x_j), \ldots, h_\alpha(x_N) \) are linearly independent on the invariant manifold \( \{e_i = 0\} \), then we also have \( d_{ij} \rightarrow c_{ij} \) when \( t \rightarrow \infty \). If the drive network is linearly coupled, namely, \( H(x_j) = Ax_j \), where \( A = (a_{ij})_{n \times n} \) is a constant matrix, then \( h_\alpha(x_j) = \sum_{j=1}^{n} a_{\alpha j} x_{ij} \). In particular, when \( A \) is a identity matrix, \( h_\alpha(x_j) \) becomes \( x_{\alpha o} \), i.e., the \( \alpha \)th variable of \( x \). This simple condition forms the basis for explaining the failure of topology identification in the case of projection synchronization of a network.

Hereinafter, we will study the topology identification problem in the drive network under three cases: complete synchronization, partial synchronization, and stability.

\textbf{A. Complete Synchronization}

Here, a network is said to be synchronized completely if the manifold of synchronized motions is stable. The manifold is defined as \( S = \{x_1 = x_2 = \cdots = x_N = s\} \), where \( s \) is the synchronous state. In this case, no element of the configuration matrix \( C \) can be identified, that is, identification of the topology fails completely. To prove this, we shall consider the effect of synchronization of the drive network on the identifying process. When the drive network [1] completely synchronizes, we have \( H(x_i) = H(x_j) = H(s) \). Thus, on the invariant manifold \( \{e_i = 0\} \), we get \( \sum_{j=1}^{N} (d_{ij} - c_{ij}) H(s) = 0 \), or equivalently, \( \sum_{j=1}^{N} d_{ij} = \sum_{j=1}^{N} c_{ij} \) (if \( C \) is a diffusive matrix satisfying zero row sum, then \( \sum_{j=1}^{N} d_{ij} = 0 \)). This means that there exist infinitely many solutions for \( d_{ij} \) and \( d_{ij} \rightarrow d_{ij}^* \neq c_{ij} \) as \( t \rightarrow \infty \).

\textbf{B. Partial Synchronization}

In this case, we consider the drive network in partial synchronization. Similarly, we define the partial synchronization of a network as follows. A network is said to be synchronized partially if some oscillators of this network synchronize with each other. Suppose that the drive network consists of \( N \) oscillators, and there are only \( m \) oscillators that completely synchronize. Without loss of generality, we set \( x_1 = x_2 = \cdots = x_m = \hat{x}, m < N \) as \( t \rightarrow \infty \), and \( \hat{x} \) is the synchronous state. In this case, only the elements of the configuration matrix \( C \) from the \( m + 1 \)th column to the \( N \)th column can successfully be identified. Since the first \( m \) oscillators are synchronized, we have \( H(x_1) = H(x_2) = \cdots = H(x_m) = H(\hat{x}) \). Thus, on the invariant manifold \( \{e_i = 0\} \), we have \( \sum_{j=1}^{m} (d_{ij} - c_{ij}) H(\hat{x}) + \sum_{j=m+1}^{N} (d_{ij} - c_{ij}) H(x_j) = 0 \). If the functions \( H(\hat{x}) \) and \( H(x_j) \), \( j = m+1, \ldots, N \), are linearly independent, then \( \sum_{j=1}^{m} (d_{ij} - c_{ij}) = 0 \), and \( d_{ij} = c_{ij} = 0 \) for \( j = m+1, \ldots, N \). That is, we get \( d_{ij} \rightarrow c_{ij}, j = m+1, \ldots, N \) as \( t \rightarrow \infty \).

In particular, when the function \( H(\cdot) \) is a linear function, the above results can be extended to the cases where the drive network is in projection synchronization, which means that all the oscillators synchronize in the sense of \( x_i = ax_j \), where \( a \) is a constant, and of a generalized synchronization such as \( x_i = bx_j + c \), in which \( b \) and \( c \) are constants.

\textbf{C. Stability}

For the last case of the drive network, we suppose that system (1) is stable, i.e., \( x_i \rightarrow x_i^* \) as \( t \rightarrow \infty \), where \( x_i^* \) is a constant vector. Then, we have \( H(x_j) \rightarrow H_i^* \) as \( t \rightarrow \infty \), in which \( H_i^* \) is a constant vector of \( n \)-dimension. Similarly, if the constant vectors \( H_1, H_2, \ldots, H_N \) are linearly independent on the synchronization manifold, then identification of topology for the drive network can be achieved. On the other hand, when the output function \( H(\cdot) \) is a 1-D state field, one can obviously find that the same linearly independent condition does not hold, and hence, identification of topology fails. Note that in an \( n \)-dimensional real space, the number of maximal linearly independent class is just \( n \), which implies that if the node number \( N \) is larger than \( n \) (the dimension of an isolated node system), then the \( n \)-dimensional vectors \( H_1, i = 1, \ldots, N \), are linearly dependent. In other words, only if the node number of a stable network is not larger than the dimension of its individual node system can the network possibly be identified. Accordingly, the stable network linked with 1-D oscillators cannot be identified.

\textbf{III. DESYNCHRONIZATION AND NUMERICAL SIMULATIONS}

For the above discussion, we can see that accurately identifying the topology of a network that is in synchronization is impossible. That is, the synchronization phenomenon in a network is an obstacle to its topology identification. Thus, the problem of topology identification is to desynchronize the network while maintaining its topology intact. To avoid synchronization in a network that needs to be identified, the following steps can be taken. 1) If \( F_i \) is unstable, then we adjust the coupling strength such that the network is desynchronized. The equation of network (1) is rewritten as

\[
\dot{x}_i = F_i(x_i) + \epsilon \sum_{j=1}^{N} c_{ij} H(x_j), \quad i = 1, 2, \ldots, N
\]

where \( \epsilon \) is a positive constant. Since the individual oscillator is unstable, it is clear that there exists a sufficiently small \( \epsilon \) such that the network (5) is desynchronized. Using the response network (2) and the adaptive-feedback control algorithm (3), we have \( d_{ij} \rightarrow c_{ij} \) as \( t \rightarrow \infty \). Then, \( c_{ij} = d_{ij}/\epsilon \) on the invariant manifold \( \{e_i = 0, i = 1, 2, \ldots, N\} \). 2) If \( F_i \) is stable, then for desynchronizing the network, we adjust the coupling term \( H(\cdot) \). For example, if the original \( H(\cdot) \) is a linear function, then we replace it with a trigonometric function or a nonlinear function. This way, although every individual oscillator is stable, the network can easily be desynchronized and becomes unstable under nonlinear coupling.
Fig. 1. Topology identification process for network (5) of Lorenz systems under network synchronization. The parameters are $\epsilon = 1$, $k_i = 1$ and $r_i = 0.5$. (a) Time evolution of valuable errors $e_{ij}$. (b) Time evolution of adaptive parameters $c_{ij}$.

Fig. 2. Topology identification process for network (5) of Lorenz systems under network desynchronization. The parameters are $\epsilon = 0.001$, $k_i = 1$, and $r_i = 0.5$. (a) Time evolution of valuable errors $e_{ij}$. (b) Time evolution of adaptive parameters $c_{ij} > 0$.

To show the effectiveness of the present method, we give an illustrative example as follows. Consider a weighted and directed linearly coupled dynamical network consisting of four identical dynamical nodes, and the topology is described by the configuration matrix $C$, where $c_{1,2} = c_{2,1} = 4$, $c_{1,3} = c_{3,1} = 7$, $c_{1,4} = c_{4,1} = 3$, $c_{2,3} = 5$, $c_{3,2} = 6$, $c_{3,4} = c_{4,3} = 2$, the other $c_{ij} = 0(i \neq j)$, and $c_{ii} = \sum_{j=1,i\neq j}^{N} c_{ij}$. Assume that the network topology is unknown before the identification process is applied. The Lorenz system, being a typical benchmark chaotic system [24], is used for illustration. In the following, we will take the Lorenz system as node dynamics to illustrate our proposed identification method.

**IV. ILLUSTRATIVE EXAMPLE**

The chaotic Lorenz system is taken as the dynamics of every node of the network: $x_i = F_i(x_i) = (\sigma(x_{i2} - x_{i1}); \gamma x_{i1} - x_{i1}x_{i3} - x_{i2}; x_{i1}x_{i2} - bx_{i3})$, for $i = 1, 2, 3, 4$ with $\sigma = 10$, $\gamma = 28$, and $b = 8/3$. For simplicity, we use (5) to describe this network model with $\epsilon = 1$, $N = 4$, and $H(x_j) = x_j$. It is easy to check that the Lorenz system satisfies the Lipschitz condition due to its boundedness [25], [26].

The synchronized behavior of the drive network for model (5) can easily be detected when the coupling strength $\epsilon$ is 1, as can be seen from Fig. 1(a). When the coupling strength $\epsilon$ becomes 0.001, the manifold of synchronized motions loses stability, and chaotic oscillators become desynchronized, as shown in Fig. 2(a). Figs. 1(b) and 2(b) show the corresponding process of identifying topology under network synchronization and desynchronization. It is clearly shown that identification fails when the network is synchronized and succeeds when the network is desynchronized.

Remarks: It is worth noting that the Lipschitz condition is important in this identification method. As a matter of fact, the Lipschitz condition may be invalid for some dynamical systems. The practicality of checking the Lipschitz condition, for some specific types of systems, depends on the boundedness property of the corresponding systems. For most nonlinear systems, however, checking the Lipschitz conditions is not a trivial task.

**V. CONCLUSION**

In this brief, we have described how a network can practically be identified by an adaptive-feedback control algorithm. The linear independence condition of the coupling terms proposed in this brief is necessary and sufficient for network identification. Synchronization is a property of a dynamical network that makes identification of the topology of the network impossible. An interesting analogy between identifying the topology of a network and identifying the parameters of a single dynamical system is worth noting here. For a stable single system,
identification of the parameters will always fail under the condition of linear dependence between the functions of inner variables. Moreover, a stable network can properly be identified if 1) the node number of the network is not larger than the dimension of the individual systems of the network, and 2) the condition of linear independence of the coupling terms holds on the manifold \( \{ e_i = 0 \} \). Intuitively, the effects of synchronization on the process of identifying the topology of network model (1) can be extended to those network models with time-delayed couplings.

REFERENCES