AIP | Chaos
An Interdisciplinary Journal of Nonlinear Science

# Reverse engineering of complex dynamical networks in the presence of time-delayed interactions based on noisy time series

Wen-Xu Wang, Jie Ren, Ying-Cheng Lai, and Baowen Li

Citation: Chaos 22, 033131 (2012); doi: 10.1063/1.4747708
View online: http://dx.doi.org/10.1063/1.4747708
View Table of Contents: http://chaos.aip.org/resource/1/CHAOEH/v22/i3
Published by the American Institute of Physics.

# Reverse engineering of complex dynamical networks in the presence of time-delayed interactions based on noisy time series

Wen-Xu Wang,[1,2] Jie Ren,[3,4,6] Ying-Cheng Lai,[2,5] and Baowen Li[3,4]
[1]Department of Systems Science, School of Management and Center for Complexity Research, Beijing Normal University, Beijing 100875, China
[2]School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, Arizona 85287, USA
[3]NUS Graduate School for Integrative Sciences and Engineering, Singapore 117456, Republic of Singapore
[4]Department of Physics and Centre for Computational Science and Engineering, National University of Singapore, Singapore 117546, Republic of Singapore
[5]Department of Physics, Arizona State University, Tempe, Arizona 85287, USA
[6]Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

(Received 10 November 2011; accepted 9 August 2012; published online 24 August 2012)

Reverse engineering of complex dynamical networks is important for a variety of fields where uncovering the full topology of unknown networks and estimating parameters characterizing the network structure and dynamical processes are of interest. We consider complex oscillator networks with *time-delayed interactions* in a noisy environment, and develop an effective method to infer the full topology of the network and evaluate the amount of time delay based solely on noise-contaminated time series. In particular, we develop an analytic theory establishing that the dynamical correlation matrix, which can be constructed purely from time series, can be manipulated to yield both the network topology and the amount of time delay simultaneously. Extensive numerical support is provided to validate the method. While our method provides a viable solution to the network inverse problem, significant difficulties, limitations, and challenges still remain, and these are discussed thoroughly. © *2012 American Institute of Physics*. [http://dx.doi.org/10.1063/1.4747708]

**Time-delayed interactions are common in complex systems arising from various fields of science and engineering. Consider, for example, a coupled oscillator network in a physical environment where noise is present. Time delay can typically occur in the node-to-node interactions. Now suppose that no prior knowledge about the nodal dynamics and the network topology is available but only a set of noise-contaminated time series can be obtained through measurements. The question is whether it is possible to deduce the full topology of the network and to estimate the amount of average time delay using the time series only. This issue belongs to the recently emerged subfield of research in complex systems: reverse engineering of complex networks (or the inverse problem). While a number of methods address the network inverse problem have appeared, to our knowledge, the issue of time-delayed interactions has not been considered. Here we present an effective method to infer the full network topology and, at the same time, to estimate the amount of average time delay in the network. In particular, we develop a physical theory to obtain a formula relating the network topology and time delay to the dynamical correlation matrix, which can be constructed purely from time series. We then show how information about the time delay encrypted in the dynamical correlation matrix can be separated from that of network topology, allowing both to be inferred in a computationally extremely efficient manner. We present numerical examples from both model and real-world complex networks to demonstrate the working of our method. Difficulties, limitations, and challenges are also discussed. Reverse-engineering of complex dynamical systems has potential applications in many disciplines, and our work represents a small step forward in this extremely challenging area.**

## I. INTRODUCTION

An outstanding and challenging problem in network science and engineering is to infer or uncover the topology and other basic characteristics of complex networks based on measured time series only. This reverse-engineering or inverse problem finds applications in a number of fields such as biomedical and technological-social sciences where complex networked systems are ubiquitous. In defense, the problem of identifying various adversarial networks based on observations is also of paramount importance. Despite previous effort in revealing the connection between network structures and dynamics,[1–5] how to infer the full topology of a complex network from dynamical behaviors is still an interesting problem, especially when knowledge about the nodal dynamics is not available.

Recent years have witnessed the emergence of a number of methods to address various aspects of the inverse problem, which include gene regulatory network inference using singular value decomposition and robust regression,[6] spike classification methods for measuring interactions among neurons from spike trains,[7] symbolically reverse engineering of coupled ordinary differential equations,[8] approaches based on response dynamics of specific oscillators,[9,10] $L_1$ norm in optimization theory,[11] noise-induced scaling laws,[12] the interplay between dynamical correlation and network structure in the presence of noise,[13] random phase resetting,[14] inner composition alignment,[15] and compressive-sensing based method.[16] Despite these efforts, the issue of time delay in the context of network reverse engineering has not been addressed. The purpose of this article is to present a theory and a purely data-driven method to uncover the network topology and to estimate the time delay at the same time, assuming there is uniform or nearly uniform time delay in the node-to-node interactions and the networked system is in a physical environment where noise is present.

Time delay is fundamental in natural systems, due to the finite propagation speed of physical signals. In addition to numerous examples in physics, situations where time delay is important include the latency times of neuronal excitations in neuroscience and finite reaction times of chemicals in chemistry. In coupled oscillator networks, the effects of time delay on dynamics under various *given* network topologies have been studied extensively.[17–21] In our case, however, the network connections, the amount of the time delay, and other properties of the network are unknown *a priori*, and our goal is to predict these by using noisy time series only. Our general point of view is that information about the network topology and time delay has been encoded in the time series from various nodes in the network. The objective of solving the inverse problem is to decode such information from time series. Our idea is that, if the networked system is in a physical environment so that the measured time series are contaminated by noise, it is possible to accomplish the task of decoding in a natural way. In particular, we construct a dynamical correlation matrix from all available time series, the elements of which are the average products of the deviations of all pairwise time series from a mean value. We shall show analytically that information about the network structure and time delay can be decoded through this matrix. In fact, as we will show in developing our theory, information about the network topology can be separated from the time delay through a generalized inverse operation of the dynamical correlation matrix, enabling a complete prediction of the underlying networked system.

To provide numerical support for our theory, we exploit three representative dynamical processes on homogeneous and heterogeneous model complex networks and on a number of real-world networks as well. We find that the presence of time delay results in a deviation in the distribution of the diagonal elements of the dynamical correlation matrix from a power law, which can be used as a preliminary criterion to determine whether there is a significant time delay in the underlying networked system. Computations reveal high accuracies in the prediction of both network topology and time delay for all combinations of dynamical processes and network models studied.

It should be emphasized that, at this stage, our theory and method apply to complex oscillator networks only, which generate continuous time series. Another key assumption that makes theoretical analysis viable is that the time delay is uniform or nearly uniform for all interactions on the network. Needless to say, this assumption is not realistic for real-world situations where the time delays associated with node-to-node interactions on the network are likely to be random and follow some sort of statistical distribution. However, insofar as the distribution is narrow, our theory is approximately valid for estimating the *average* time delay on the network. At the present, we are unable to extend our theory to situations where the time-delay distributions are broad. We suspect that, in more general situations, some drastically different approach may be needed, warranting further effort in this direction.

In Sec. II, we develop our theory, which leads to a computationally extremely efficient, completely data-driven method to uncover the network topology and to estimate the time delay. In Sec. III, we present results from extensive numerical computations to verify our theory and validate the corresponding method. In Sec. IV, we consider a variant of the oscillator network model and demonstrate that a theory can be developed which leads essentially the same method to address the oscillator-network inverse problem. Conclusion and discussion of the limitations of our theory and method are presented in Sec. V.

## II. THEORY AND PREDICTION METHOD

We present our theory and method by considering a network of $N$ coupled oscillators. Each oscillator, when decoupled, satisfies $\dot{\vec{x}}_i = \mathbf{F}_i[\vec{x}_i]$, where $\vec{x}_i$ denotes the $d$-dimensional state variable of node $i$. The dynamics of the whole time-delayed system in a noisy environment can be described as

$$\dot{\vec{x}}_i(t) = \mathbf{F}_i[\vec{x}_i(t)] - c\sum_{j=1}^{N} L_{ij}\mathbf{H}[\vec{x}_j(t-\tau)] + \vec{\eta}_i(t), \qquad (1)$$

where $c$ is the coupling strength, $\mathbf{H}$ denotes the coupling function, and $L_{ij}$ is the element of the Laplacian matrix characterizing the topology of the underlying network, which satisfies $L_{ij} = -1$ if $j$ connects to $i$ (otherwise 0) for $i \neq j$, and $L_{ii} = k_i$, where $k_i$ is the degree of node $i$. The quantity $\tau$ denotes the uniform time delay on the network, and $\vec{\eta}_i$ is a $d$-dimensional stochastic process representing noise on node $i$. For convenience, in the following we use the upper arrow sign to denote the $d$-dimensional state variable. The standard procedure of linearization[1,17] can be carried out by letting $\vec{x}_i = \bar{x}_i + \vec{\xi}_i$, where $\bar{x}_i$ is the counterpart of $x_i$ in the absence of noise. The $d$-dimensional dynamical process governing the fluctuations associated with the dynamical process on the $i$th oscillator is governed by the following variational equation:

$$\dot{\vec{\xi}}_i(t) = D\mathbf{F}_i \cdot \vec{\xi}_i(t) - c\sum_{j=1}^{N} L_{ij} D\mathbf{H} \cdot \vec{\xi}_j(t-\tau) + \vec{\eta}_i(t), \qquad (2)$$

where $D\mathbf{F}_i$ and $D\mathbf{H}$ denote the $d \times d$ Jacobian matrices of the intrinsic dynamics $\mathbf{F}_i$ and the coupling function $\mathbf{H}$, respectively. Decomposing Eq. (2) in terms of the eigenmodes, we obtain

$$\dot{\vec{\epsilon}}_\alpha(t) = \sum_\beta D\mathbb{F}_{\alpha\beta} \cdot \vec{\epsilon}_\beta(t) - c\lambda_\alpha D\mathbf{H} \cdot \vec{\epsilon}_\alpha(t-\tau) + \vec{\zeta}_\alpha(t), \qquad (3)$$

where instead of the index $i,j$ running on the real space of networks, the indices $\alpha$ and $\beta$ run in the eigenspace. The various quantities in Eq. (3) are given by

$$\vec{\epsilon}_\alpha = \sum_i \psi_{\alpha i}\vec{\xi}_i,$$
$$\vec{\zeta}_\alpha = \sum_i \psi_{\alpha i}\vec{\eta}_i,$$
$$D\mathbb{F}_{\alpha\beta} = \sum_i \psi_{\alpha i} D\mathbf{F}_i \psi_{\beta i},$$

where $\psi_{\alpha j}$ denotes the $\alpha$th normalized eigenvector of the Laplacian matrix, and $\lambda_\alpha$ are the corresponding eigenvalues that satisfy $0 = \lambda_0 < \lambda_1 \leq \cdots \leq \lambda_{N-1}$. Under the approximation $D\mathbf{F}_i \approx D\mathbf{F}$ so that

$$D\mathbb{F}_{\alpha\beta} = D\mathbf{F}\delta_{\alpha\beta},$$

Eq. (3) can be reduced to

$$\dot{\vec{\epsilon}}_\alpha(t) = D\mathbf{F}\cdot\vec{\epsilon}_\alpha(t) - c\lambda_\alpha D\mathbf{H}\cdot\vec{\epsilon}_\alpha(t-\tau) + \vec{\zeta}_\alpha(t). \quad (4)$$

From the covariance of Gaussian noise

$$\langle\vec{\eta}_i(t)\vec{\eta}_j^T(t')\rangle = \sigma^2\mathbf{I}_d\delta_{ij}\delta(t-t'),$$

with $\mathbf{I}_d$ being the $d$-dimensional identity matrix and $\sigma^2$ the noise strength, we obtain

$$\langle\vec{\zeta}_\alpha(t)\vec{\zeta}_\beta^T(t')\rangle = \sigma^2\mathbf{I}_d\delta_{\alpha\beta}\delta(t-t'), \quad (5)$$

indicating that the stochastic process mapped into the eigenspace still represents Gaussian noise. Assuming small time delay, we can apply the first-order approximation

$$\vec{\epsilon}_\alpha(t-\tau) = \vec{\epsilon}_\alpha(t) - \tau\dot{\vec{\epsilon}}_\alpha(t), \quad (6)$$

which yields

$$(\mathbf{I}_d - c\tau\lambda_\alpha D\mathbf{H})\dot{\vec{\epsilon}}_\alpha(t) = -(c\lambda_\alpha D\mathbf{H} - D\mathbf{F})\vec{\epsilon}_\alpha(t) + \vec{\zeta}_\alpha(t). \quad (7)$$

Letting

$$\mathbf{B} \equiv (\mathbf{I}_d - c\tau\lambda_\alpha D\mathbf{H})^{-1},$$
$$\mathbf{A} \equiv \mathbf{B}(c\lambda_\alpha D\mathbf{H} - D\mathbf{F}),$$

and following standard stochastic calculus,[22] we get the solution

$$\vec{\epsilon}_\alpha(t) = e^{-\mathbf{A}t}\vec{\epsilon}_\alpha(0) + \int_0^t e^{-\mathbf{A}(t-t')}\mathbf{B}\vec{\zeta}_\alpha(t')dt'. \quad (8)$$

Since we are interested in the regime where oscillator states are perturbed from the synchronized manifold by noise, it is reasonable to assume that the state variables of the system will not diverge. Thus, in the long time limit, the initial-condition term can be neglected,[13] which yields

$$\mathbf{A}\langle\vec{\epsilon}_\alpha\vec{\epsilon}_\alpha^T\rangle + \langle\vec{\epsilon}_\alpha\vec{\epsilon}_\alpha^T\rangle\mathbf{A} = \sigma^2\mathbf{B}\mathbf{B}^T. \quad (9)$$

The general solution of $\langle\vec{\epsilon}_\alpha\vec{\epsilon}_\alpha^T\rangle$, a $d\times d$ covariance matrix about the $d$-dimensional states of the $\alpha$th oscillator in the eigenspace, can be written as[23]

$$vec(\langle\vec{\epsilon}_\alpha\vec{\epsilon}_\alpha^T\rangle) = \sigma^2 vec(\mathbf{B}\mathbf{B}^T)/(\mathbf{I}_d\otimes\mathbf{A} + \mathbf{A}\otimes\mathbf{I}_d), \quad (10)$$

where the operator $vec(\mathbf{X})$ creates a column vector from a matrix $\mathbf{X}$ by stacking the columns of $\mathbf{X}$ below one another.

It is difficult to apply the solution in Eq. (10) to general networked systems with arbitrary (but unknown) coupling functions. To obtain a practically useful formula, further simplifications and approximations are needed. In the following, we assume one-dimensional state variable so that $D\mathbf{H} = 1$. The vector notation can then be dropped. Equation (10) becomes

$$\langle\epsilon_\alpha^2\rangle = \frac{\sigma^2}{2c}\frac{1}{(1-c\tau\lambda_\alpha)(\lambda_\alpha - D\mathbf{F}/c)}. \quad (11)$$

Returning to real variables from the eigenspace by inserting

$$\xi_i = \sum_\alpha \psi_{\alpha i}\epsilon_\alpha,$$

into the correlation function $C_{ij} = \langle\xi_i\xi_j\rangle$ in the real space between any two nodes, we have

$$C_{ij} = \sum_{\alpha=1}^{N-1}\psi_{\alpha i}\psi_{\alpha j}\langle\epsilon_\alpha^2\rangle,$$

and, consequently,

$$C_{ij} = \frac{\sigma^2}{2c}\sum_{\alpha=1}^{N-1}\frac{\psi_{\alpha i}\psi_{\alpha j}}{(1-c\tau\lambda_\alpha)(\lambda_\alpha - D\mathbf{F}/c)}. \quad (12)$$

Under the approximation that the term $D\mathbf{F}/c$ can be neglected and the time delay $\tau$ is small, Eq. (12) for the dynamical correlation can be expanded as

$$C_{ij} \approx \frac{\sigma^2}{2c}\sum_{\alpha=1}^{N-1}\frac{1+c\tau\lambda_\alpha}{\lambda_\alpha}\psi_{\alpha i}\psi_{\alpha j} = \frac{\sigma^2}{2c}[\mathbf{L}^\dagger + c\tau\mathbf{I}_N]_{ij}, \quad (13)$$

wherein under the influence of noise of variance $\sigma^2$, the dynamical correlation matrix $\mathbf{C}$ is connected explicitly with the time delay $\tau$ and the structure information in terms of the matrix

$$\mathbf{L}^\dagger = \sum_{\alpha=1}^{N-1}\psi_{\alpha i}\psi_{\alpha j}/\lambda_\alpha, \quad (14)$$

which is the pseudo-inverse of the Laplacian matrix. Note that the time delay has no effect on the cross-correlation elements except the auto-correlations due to the identity matrix. Following Eq. (13), the diagonal elements $C_{ii}$ of the dynamical correlation matrix can be obtained by expanding $\mathbf{L}^\dagger$ in terms of the underlying network structural properties[13]

$$C_{ii} \approx \frac{\sigma^2}{2c}[\mathbf{K}^{-1} + \mathbf{K}^{-1}\mathbf{P}\mathbf{K}^{-1} + \mathbf{K}^{-1}\mathbf{P}\mathbf{K}^{-1}\mathbf{P}\mathbf{K}^{-1}]_{ii} + \frac{\sigma^2\tau}{2}$$
$$\approx \frac{\sigma^2}{2ck_i}\left(1 + \frac{1}{\langle k\rangle}\right) + \frac{\sigma^2\tau}{2}, \quad (15)$$

where $\mathbf{K} = \mathrm{diag}(k_1, \cdots, k_N)$ is the degree matrix, $\mathbf{P}$ is the adjacency matrix such that $\mathbf{L} = \mathbf{K} - \mathbf{P}$, and $\langle k \rangle$ denotes the average degree. We see that the fluctuations $C_{ii}$ at node $i$ depend both on its local structure $k_i$ and the time delay $\tau$. For $\tau = 0$, this result is consistent with the recently discovered noise-induced scaling law,[12] derived there by a power-spectral analysis.

The off-diagonal elements of $\mathbf{L}$ contain complete information about the network structure while its diagonal elements can be obtained from the off-diagonal ones. We thus focus on the off-diagonal elements. For $i \neq j$, following Eq. (12), we obtain the generalized inverse matrix $\mathbf{C}^\dagger$ as

$$C^\dagger_{ij} \approx \frac{2c}{\sigma^2} \sum_{\alpha=1}^{N-1} \lambda_\alpha (1 - c\tau\lambda_\alpha) \psi_{\alpha i} \psi_{\alpha j} = \frac{2c}{\sigma^2} [\mathbf{L} - c\tau \mathbf{L}^2]_{ij}. \quad (16)$$

Using the relation $\mathbf{L} = \mathbf{K} - \mathbf{P}$, we can cast Eq. (16) in the following form:

$$C^\dagger_{ij} = \frac{2c}{\sigma^2} [\mathbf{L} + c\tau(\mathbf{KP} + \mathbf{PK} - \mathbf{K}^2 - \mathbf{P}^2)]_{ij}. \quad (17)$$

To the off-diagonal elements $(i, j)$, the diagonal matrix $\mathbf{K}^2$ has no contributions and $\mathbf{P}^2$ contributes the value $l_{ij}$, the number of two-step paths connecting $i$ with $j$. If the network has a high average degree, the contribution of $l_{ij}$ can be neglected. We thus have

$$\frac{\sigma^2}{2c} C^\dagger_{ij} \approx \begin{cases} L_{ij} + c\tau(k_i + k_j), & \text{if } i \text{ connects with } j \\ 0. & \text{otherwise} \end{cases}, \quad (18)$$

Equation (18) is one of our main results for time-series based solution to the network inverse problem, which indicates that the full topology of the network can be inferred through the off-diagonal elements $C^\dagger_{ij}$ of the dynamical correlation matrix based solely on measured time series.

Once $\mathbf{L}$ has been predicted, the time delay $\tau$ can be estimated, e.g., from Eq. (16). We obtain

$$\tau \approx \left\langle \frac{[\mathbf{L} - \frac{\sigma^2}{2c}\mathbf{C}^\dagger]_{ij}}{c[\mathbf{L}^2]_{ij}} \right\rangle_{i \neq j, L_{ij} \neq 0, (\mathbf{L}^2)_{ij} \neq 0}, \quad (19)$$

here the subscript in the average $\langle \cdot \rangle$ covers all possible pairs of $i$ and $j$ by excluding the diagonal elements in the matrices $\mathbf{L}$ and $\mathbf{L}^2$, and all pairs with zero elements in the matrix $\mathbf{L}$ or $\mathbf{L}^2$. Excluding zero elements can effectively reduce the estimation error for $\tau$.

## III. NUMERICAL RESULTS

The main theoretical results, Eqs. (18) and (19), are derived under a number of assumptions and approximations. To lend credence to their validity for complex oscillator networks, we apply the results to a number of model and real-world networks in the presence of noise and uniform time delay. For each network, we implement three dynamical processes: (i) Consensus dynamics:[24]

$$\dot{x}_i(t) = c \sum_{j=1}^{N} P_{ij}[x_j(t-\tau) - x_i(t-\tau)] + \eta_i,$$

(ii) Rössler dynamics:[25]

$$\begin{cases} \dot{x}_i = -y_i - z_i + c\sum_{j=1}^{N} P_{ij}[x_j(t-\tau) - x_i(t-\tau)] + \eta_i, \\ \dot{y}_i = x_i + 0.2y_i + c\sum_{j=1}^{N} P_{ij}(y_j - y_i), \\ \dot{z}_i = 0.2 + z_i(x_i - 9.0) + c\sum_{j=1}^{N} P_{ij}(z_j - z_i), \end{cases}$$

and (iii) Kuramoto phase dynamics:[26]

$$\dot{\theta}_i(t) = \omega_i + c\sum_{j=1}^{N} \sin[\theta_j(t-\tau) - \theta_i(t-\tau)] + \eta_i,$$

where $\theta_i$ and $\omega_i$ are the phase and the natural frequency of oscillator $i$, respectively.

Time series are collected from all nodes. The element of the dynamical correlation matrix between two arbitrary nodes $i$ and $j$ is calculated as

$$C_{ij} = \langle [x_i(t) - \bar{x}(t)] \cdot [x_j(t) - \bar{x}(t)] \rangle_t,$$

where $\bar{x}(t) = (1/N)\sum_{i=1}^{N} x_i(t)$ and $\langle \cdot \rangle_t$ denotes long-time average. For the Rössler dynamics, $x_i(t)$ stands for the $x$ component of the $i$th oscillator and, for the Kuramoto dynamics, $x_i(t)$ stands for the phase variable $\theta_i(t)$ of the $i$th oscillator.

Figure 1 shows an example of the dependence of fluctuations $C_{ii}$ on the time delay for three dynamical processes on heterogeneous (scale-free) and homogeneous (random) networks. The results are in good agreement with the theoretical prediction from Eq. (15), except a deviation for the case of Rössler dynamics, which is mainly caused by the simplification of one-dimensional variable for obtaining Eq. (15), as the Rössler dynamics is intrinsically three-dimensional. Note that, in the absence of time delay, the dependence of $C_{ii}$ on the node degree $k_i$ can be described as a power law:[12,13] $C_{ii} \sim k_i^{-1}$, regardless of the specific dynamical processes on the network. For $\tau \neq 0$, deviation of $C_{ii}$ from the power-law behavior can then be used to assess, *preliminarily*, whether there is a significant time delay in the underlying networked system: a more severe deviation suggests a larger value of the time delay.

Having calculated the dynamical correlation matrix $\mathbf{C}$, we can infer the details of the network connections through Eq. (18) via the generalized inverse of $\mathbf{C}$. Figure 2 shows the distribution of the off-diagonal elements of $[\sigma^2/(2c)]\mathbf{C}^\dagger$. We observe a bimodal behavior with two peaks: one centered at a negative value which corresponds to existent links, and another centered at zero which indicates non-existent links. Without time delay, the hump in the distribution for the existent links should be centered at $-1$. In the presence of time delay, due to the contribution of the term $c\tau(k_i + k_j) \approx 2c\tau\langle k \rangle$, the center of the hump

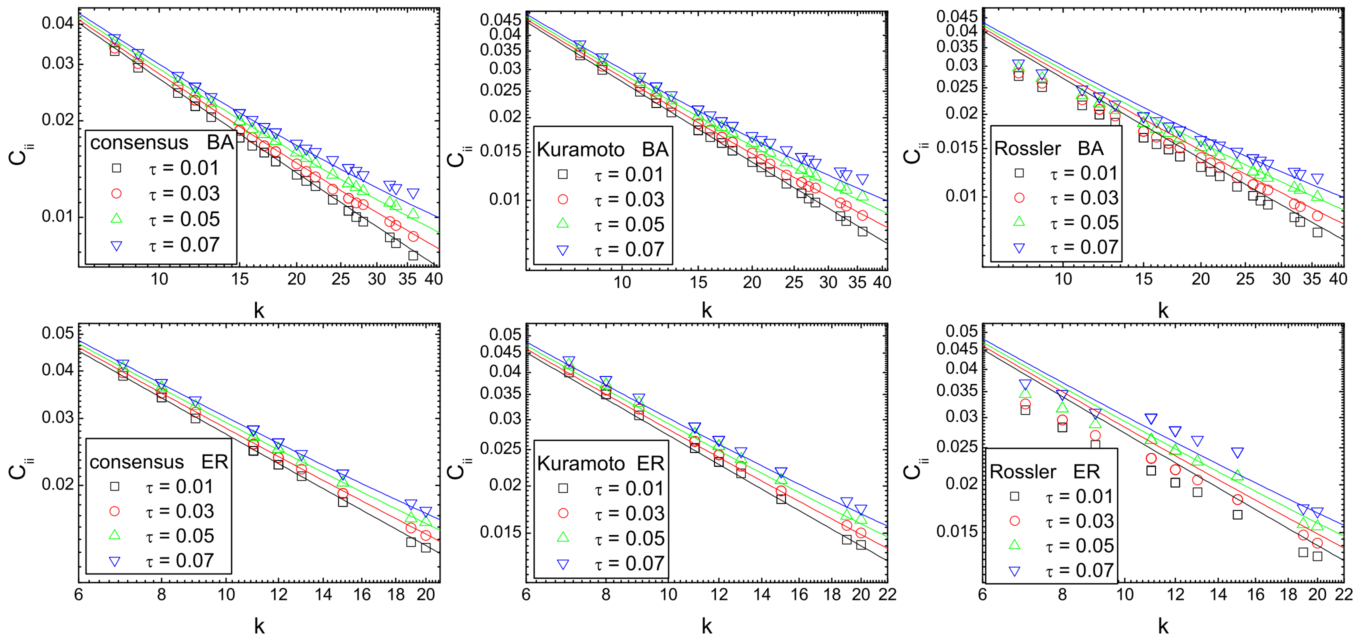


FIG. 1. Diagonal elements $C_{ii}$ of the dynamical correlation matrix as a function of node degree $k$ for three dynamical processes with different values of the time delay $\tau$ on scale-free and random networks. Square, circle, triangle and reverse triangle denote $\tau = 0.01, 0.05, 0.07$, and $0.09$, respectively. The curves are the theoretical prediction from Eq. (15). The sizes of model networks are 100 and the average degree is 10. The noise strength $\sigma^2$ is 0.1 and the coupling strength $c$ is 0.2.

will shift toward zero. The larger the time delay, the more significant the shift will be. For example, as shown in Fig. 2, the amount of the shift in the negative peak is $2c\tau\langle k\rangle = 0.2$. Since $c = 0.2$ and $\langle k\rangle = 10$ in the example, we obtain $\tau = 0.05$, which agrees with the pre-assumed value of the time delay very well. To separate the two humps, a threshold is needed, where all existent links in the network are identified by the elements in $\mathbf{C}^{\dagger}$ that lie below the threshold. In particular, the subscript $ij$ of a chosen element below the threshold indicates a link between nodes $i$ and $j$. We set the threshold at the minimum value of the fitted curve between two humps.

It is noteworthy that the bimodal distribution in Fig. 2 will be affected by the node degree distribution and the time delay. In particular, when the fluctuation of the degree distribution is large, e.g., power-law distribution, and the time delay is large, the gap between two humps may shrink and even vanish, rendering the overlap of them. In this case, the network structure and time delay might not be accurately predicted due to the difficulty in separating two humps. Details will be presented below with respect to different degree distributions and time delays.

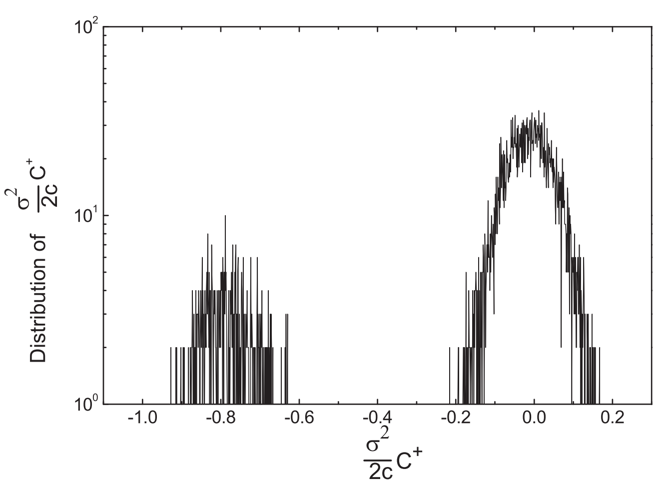


FIG. 2. Example of the distribution of the values of elements of the generalized inverse $\mathbf{C}^{\dagger}$ of the dynamical correlation matrix $\mathbf{C}$ for consensus dynamics associated with a scale-free network, where $\tau = 0.05$. The bimodal behavior is present for Kuramoto model and Rössler dynamics as well.

The performance of our prediction method can be characterized by the success rate $S_e$ of the existent links, which is defined as the ratio of the number of successfully predicted existent links to the total number of existent ones. As shown in Figs. 3(a)–3(c), our method yields high success rates for different values of the time delay $\tau$, regardless of the nodal dynamics and of the network structure.

After the network topology $\mathbf{L}$ is predicted, we can estimate the time delay $\tau$ through Eq. (19). As demonstrated in Fig. 4, the predicted values of $\tau$ are quite close to the real values for almost all dynamical processes and network structures considered. The deviation of numerical results from theory in the case of Kuramoto dynamics is attributed to the sinusoidal coupling function, for which the linear coupling function assumed in our theory is only a crude approximation.

We also test the applicability of our method for situations where the time delays on the network are not uniform. In this case, we expect our method to give the *average* time delay if it is not too large and if the statistical distribution of the time delays is not broad. Specifically, we consider a random network under consensus dynamics with time delays randomly distributed within a certain range. The success rate $S_e$ as a function of the average time delay $\tau$ among all pairs of nodes for different ranges of time delay is shown in Fig. 5(a). We obtain high success rate. Figure 5(b) shows the predicted average time delay $\tau'$ versus the original time delay $\tau$ for different ranges of time delay. The predicted $\tau'$ is in good

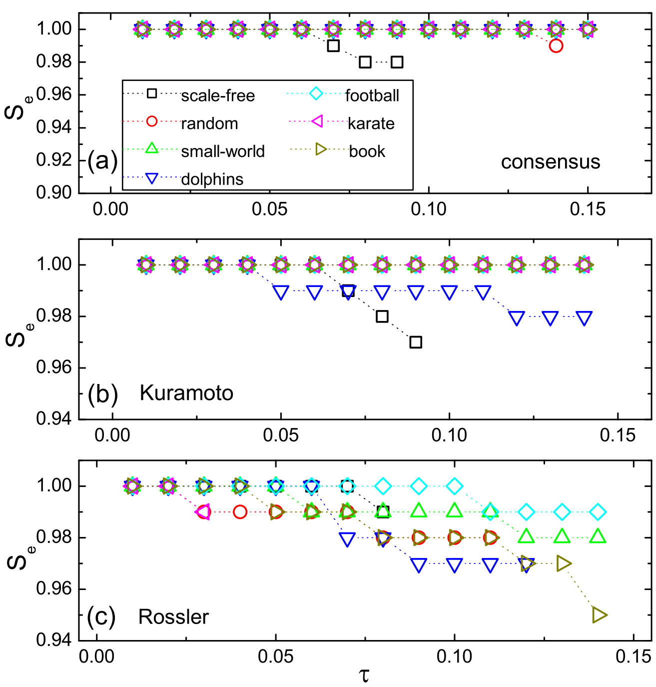


FIG. 3. Success rate of prediction of existent links $S_e$ for (a) consensus dynamics, (b) Kuramoto oscillators, and (c) Rössler dynamics as a function of time delay $\tau$ for a number of model and real-world networks: scale-free networks (scale-free),[27] random network (random),[28] small-world network (small-world),[29] dolphin social network (dolphins),[30] network of American football games among colleges (football),[31] friendship network of karate club (karate),[32] and network of political book purchases (book).[33] Other parameters are the same as in Fig. 1. The success rate of nonexistent links is higher than 0.99 for all considered cases and thus are not shown.

agreement with $\tau$. These results demonstrate that our approach is applicable even under reasonably constrained, inhomogeneous time delays on coupled oscillators networks.

## IV. VARIANT OF COUPLED OSCILLATOR NETWORK MODEL

While our theory and the prediction method are based on the system model Eq. (1), a similar theory can be developed for variants of the model. For instance, one can consider the following system:[24]

$$\dot{x}_i(t) = \mathrm{F}_i[x_i(t)] - c\sum_{j=1}^{N} P_{ij}[x_i(t) - x_j(t-\tau)] + \eta_i(t), \quad (20)$$

where the dynamics of each isolated oscillator $i$ is described by $\dot{x}_i = \mathrm{F}_i(x_i)$. Again, we consider one-dimensional state variable. The quantity $P_{ij}$ denotes the element of the adjacency matrix of the underlying network, $c$ denotes coupling strength, $\tau$ is the time delay which occurs only for the state information transmitted from its connected neighbors other than the dynamics of itself, and $\eta_i$ denotes noise. We linearize Eq. (20) by using $x_i = \bar{x}_i + \epsilon_i$, where $\bar{x}_i$ is the counterpart of $x$ in the absence of noise. The dynamical equation describing fluctuations becomes

$$\dot{\epsilon}_i(t) = (D\mathrm{F}_i - ck_i)\cdot\epsilon_i(t) + c\sum_{j=1}^{N} P_{ij}\cdot\epsilon_j(t-\tau) + \eta_i(t), \quad (21)$$

where $D\mathrm{F}_i$ denotes the derivative of function $\mathrm{F}_i$ with respect to $x_i$. Decomposing Eq. (21) in terms of the normal modes, we have

$$\dot{\xi}_\alpha(t) = \sum_\beta \widehat{D\mathrm{F}}_{\alpha\beta}\cdot\xi_\beta(t) + c\lambda_\alpha\xi_\alpha(t-\tau) + \zeta_\alpha(t), \quad (22)$$

where $\xi_\alpha = \sum_i \psi_{\alpha i}\epsilon_i$, $\zeta_\alpha = \sum_i \psi_{\alpha i}\eta_i$,

$$\widehat{D\mathrm{F}}_{\alpha\beta} = \sum_i \psi_{\alpha i}(D\mathrm{F}_i - ck_i)\psi_{\beta i},$$

$\psi_{\alpha j}$ denotes the $\alpha$th normalized eigenvector of the adjacency matrix, and $\lambda_\alpha$ are the corresponding eigenvalues for $\alpha = 0, \cdots, N-1$. Using the approximation $D\mathrm{F}_i \approx 0$ and $k_i \approx \langle k\rangle$ so that

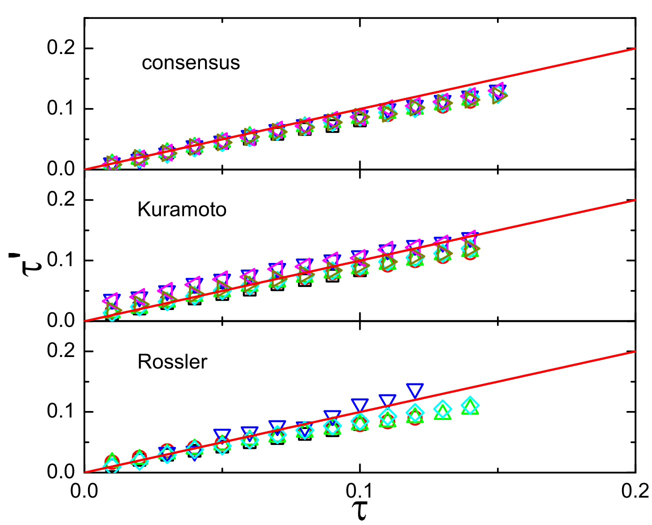


FIG. 4. Predicted time delay $\tau'$ from Eq. (19) versus the true (pre-assumed) values for the three dynamical processes on a number of model and real-world networks. The symbols denote the same networks as in Fig. 3. The lines are $\tau' = \tau$. Other parameters are the same as in Fig. 1.

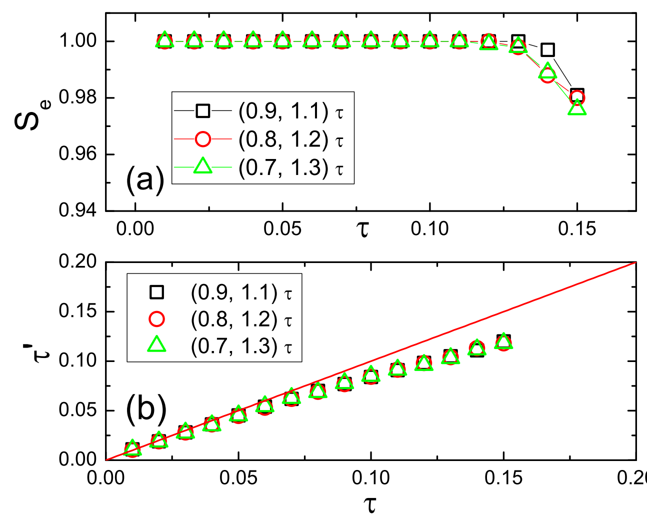


FIG. 5. (a) Success rate of prediction of existent links $S_e$ as a function of the average time delay $\tau$ for different ranges of time delays for random consensus networks. (b) Predicted average time delay $\tau'$ versus the original time delay $\tau$ for different ranges of time delay. The lines are $\tau' = \tau$. Other parameters are the same as in Fig. 1.

$$\widehat{DF}_{\alpha\beta} = -c\langle k\rangle\delta_{\alpha\beta},$$

we can write Eq. (22) as

$$\dot{\xi}_\alpha(t) = -c\langle k\rangle\xi_\alpha(t) + c\lambda_\alpha\xi_\alpha(t-\tau) + \zeta_\alpha(t). \tag{23}$$

Since the covariance of Gaussian noise is given by $\langle\eta_i(t)\eta_i(t')\rangle = \sigma^2\delta_{ij}\delta(t-t')$, we can obtain a similar relation:

$$\langle\zeta_\alpha(t)\zeta_\beta(t')\rangle = \sigma^2\delta_{\alpha\beta}\delta(t-t').$$

Applying the first-order approximation $\xi(t-\tau) = \xi(t) - \tau\dot{\xi}(t)$ for small time delay $\tau$, we obtain

$$(1 + c\tau\lambda_\alpha)\dot{\xi}_\alpha(t) = -c(\langle k\rangle - \lambda_\alpha)\xi_\alpha(t) + \zeta_\alpha(t). \tag{24}$$

In the long time limit $t \to \infty$, Eq. (24) has the solution

$$\xi_\alpha(t) = \frac{1}{1 + c\tau\lambda_\alpha}\int_0^t \exp\left[-\frac{c(\langle k\rangle - \lambda_\alpha)(t-t')}{1 + c\tau\lambda_\alpha}\right]\zeta_\alpha(t')dt, \tag{25}$$

which yields

$$\begin{aligned}\langle\xi_\alpha^2\rangle &= \frac{1}{(1 + c\tau\lambda_\alpha)^2}\int_0^t\int_0^t \exp\left[-\frac{c(\langle k\rangle - \lambda_\alpha)(t-t')}{1 + c\tau\lambda_\alpha}\right] \\ &\quad \exp\left[-\frac{c(\langle k\rangle - \lambda_\alpha)(t-t'')}{1 + c\tau\lambda_\alpha}\right]\langle\zeta_\alpha(t')\zeta_\alpha(t'')\rangle dt'dt'' \\ &= \frac{\sigma^2}{(1 + c\tau\lambda_\alpha)^2}\int_0^t \exp\left[-\frac{2c(\langle k\rangle - \lambda_\alpha)(t-t')}{1 + c\tau\lambda_\alpha}\right]dt' \\ &= \frac{\sigma^2}{2c(\langle k\rangle - \lambda_\alpha)(1 + c\tau\lambda_\alpha)}.\end{aligned} \tag{26}$$

Now turn to the fluctuations $\epsilon_i$. Substituting $\epsilon_i = \sum_\alpha \psi_{\alpha i}\xi_\alpha$ into the correlation function $C_{ij} = \langle\epsilon_i\epsilon_j\rangle$ between any two nodes, we have

$$\begin{aligned}C_{ij} &= \sum_{\alpha=0}^{N-1}\psi_{\alpha i}\psi_{\alpha j}\langle\xi_\alpha^2\rangle \\ &= \frac{\sigma^2}{2c}\sum_{\alpha=0}^{N-1}\frac{\psi_{\alpha i}\psi_{\alpha j}}{(\langle k\rangle - \lambda_\alpha)(1 + c\tau\lambda_\alpha)}.\end{aligned} \tag{27}$$

The off-diagonal elements of $\mathbf{C}$ contain complete information about the network structure while its diagonal elements can be obtained from off-diagonal ones. It thus suffices to consider off-diagonal elements only. For $i \neq j$, Eq. (27) gives the generalized inverse matrix $\mathbf{C}^\dagger$ as

$$\begin{aligned}C_{ij}^\dagger &= \frac{2c}{\sigma^2}\sum_{\alpha=0}^{N-1}(\langle k\rangle - \lambda_\alpha)(1 + c\tau\lambda_\alpha)\psi_{\alpha i}\psi_{\alpha j} \\ &\approx \frac{2c}{\sigma^2}\left[(c\tau\langle k\rangle - 1)\mathbf{P} - c\tau\mathbf{P}^2\right]_{ij}.\end{aligned} \tag{28}$$

After $\mathbf{P}$ is successfully predicted, the time delay $\tau$ can be estimated from Eq. (28) as

$$\tau \approx \left\langle \frac{\left[\mathbf{P} + \frac{\sigma^2}{2c}\mathbf{C}^\dagger\right]_{ij}}{c\left[\langle k\rangle\mathbf{P} - \mathbf{P}^2\right]_{ij}} \right\rangle_{i\neq j, P_{ij}\neq 0, (\mathbf{P}^2)_{ij}\neq 0}, \tag{29}$$

where the average $\langle\cdot\rangle$ covers all possible $i$ and $j$ by excluding the diagonal elements in matrices $\mathbf{P}$ and $\mathbf{P}^2$, and all zero off-diagonal elements in $\mathbf{P}$ and $\mathbf{P}^2$. Excluding these zero elements can effectively reduce the error in estimating $\tau$.

## V. CONCLUSION AND OUTLOOK

In summary, we have developed a physical theory to address the inverse problem for complex oscillator networked systems in the presence of time delay and noise, based solely on measured time series. Especially, we have obtained a formula relating the generalized inverse of the dynamical correlation matrix, which can be computed purely from data, to the structural Laplacian (or adjacency) matrix and the amount of time delay. Under reasonable approximations, the network topology and the effect of time delay can be separated, leading to a computationally extremely efficient method for inferring the network topology and for estimating the time delay. The validity of the method has been tested numerically using a variety of combinations of nodal dynamics and network topology, including a number of real-world network structures. Our method is completely data driven, and we expect it to be applicable to reverse engineering of complex networks in a variety of fields, such as biomedical and social sciences where such systems are ubiquitous.

It is noteworthy that our theory and method are valid under a number of assumptions/approximations: (i) oscillator networks that generate continuous time series, (ii) approximately one-dimensional nodal dynamics, (iii) uniform or nearly uniform time delays in all node-to-node interactions on the network, (iv) the system's remaining in a steady state, (v) symmetric interactions between nodes, i.e., no directionality. These assumptions/approximations are necessary for the development of the theory and the resulting prediction method. For instance, the standard linearization can be implemented only when the system is in a steady state; otherwise if the time series is obtained from some transient states, the linearization is no longer valid. The assumption of uniform or nearly uniform time delays is needed for decomposing Eq. (2) in terms of the eigenmodes. In addition, our theory and method at this stage are not applicable to weighted networks with inhomogeneous interaction strengths because of the approximation used in our derivations. In particular, the disagreement between the theory and simulations resulting from the approximations will lead to inaccuracy in the predicted weights and a considerable overlap of bimodal distribution in such networks. The overlap renders the difficulty in necessary separation of existent links from zero elements in the Laplacian matrix, so that both the link weights and the network structure inferred by the current method would not be reliable. In this regard, solving the "inverse" problem in directed, weighted dynamical networks with non-uniform time delays remains to be a challenging task. Our method for treating identical nodal dynamics might be possibly extended to nonidentical cases with small parameter mismatch in the framework of linearized equations, as has been explored in synchronization of coupled nonidentical oscillators.[34,35] We hope our work can stimulate further efforts in articulating more powerful approaches to the inverse problem in more general networked systems beyond the simplified cases studied in the present work.

## ACKNOWLEDGMENTS

J.R. thanks Dr. Gang Yan for useful discussions. W.X.W. and Y.C.L. are supported by AFOSR under Grant No. FA9550-10-1-0083. W.X.W. is supported by NSFC under Grant No. 11105011.

[1] L. M. Pecora and T. L. Carroll, Phys. Rev. Lett. **80**, 2109 (1998); M. Barahona and L. M. Pecora, Phys. Rev. Lett. **89**, 054101 (2002).
[2] M. Boguñá and R. Pastor-Satorras, Phys. Rev. E **66**, 047104 (2002).
[3] T. Nishikawa, A. E. Motter, Y.-C. Lai, and F. C. Hoppensteadt, Phys. Rev. Lett. **91**, 014101 (2003).
[4] J. G. Restrepo, E. Ott, and B. R. Hunt, Phys. Rev. Lett. **97**, 094102 (2006).
[5] J. Ren and B. Li, Phys. Rev. E **79**, 051922 (2009).
[6] M. K. S. Yeung, J. Tegneŕ, and J. J. Collins, Proc. Natl. Acad. Sci. U.S.A. **99**, 6163 (2002).
[7] R. Gütig, A. Aertsen, and S. Rotter, Neural Comput. **14**, 121 (2002); G. Pipa and S. Grün, Neurocomputing **52**, 31 (2003).
[8] J. Bongard and H. Lipson, Proc. Natl. Acad. Sci. U.S.A. **104**, 9943 (2007).
[9] M. Timme, Phys. Rev. Lett. **98**, 224101 (2007).
[10] S. G. Shandilya and M. Timme, New J. Phys. **13**, 013004 (2011).
[11] D. Napoletani and T. D. Sauer, Phys. Rev. E **77**, 026103 (2008).
[12] W.-X. Wang, Q. Chen, L. Huang, Y.-C. Lai, and M. A. F. Harrison, Phys. Rev. E **80**, 016116 (2009).
[13] J. Ren, W.-X. Wang, B. Li, and Y.-C. Lai, Phys. Rev. Lett. **104**, 058701 (2010).
[14] Z. Levnajić and A. Pikovsky, Phys. Rev. Lett. **107**, 034101 (2011).
[15] S. Hempel, A. Koseska, J. Kurths, and Z. Nikoloski, Phys. Rev. Lett. **107**, 054101 (2011).
[16] W.-X. Wang, R. Yang, Y.-C. Lai, V. Kovanis, and M. A. F. Harrison, Europhys. Lett. **94**, 48006 (2011); W.-X. Wang, R. Yang, Y.-C. Lai, V. Kovanis, and C. Grebogi, Phys. Rev. Lett. **106**, 154101 (2011).
[17] M. Dhamala, V. K. Jirsa, and M. Ding, Phys. Rev. Lett. **92**, 074104 (2004).
[18] W. Kinzel, A. Englert, G. Reents, M. Zigzag, and I. Kanter, Phys. Rev. E **79**, 056207 (2009).
[19] G. Yan, G. Chen, J. Lu, and Z. Q. Fu, Phys. Rev. E **80**, 056116 (2009).
[20] V. Flunkert, S. Yanchuk, T. Dahms, and E. Schöll, Phys. Rev. Lett. **105**, 254101 (2010).
[21] D. Hunt, G. Korniss, and B. K. Szymanski, Phys. Rev. Lett. **105**, 068701 (2010); S. Hod, Phys. Rev. Lett. **105**, 208701 (2010).
[22] C. W. Gardiner, *Handbook of Stochastic Methods*, 2nd ed. (Springer, 1997).
[23] R. A. Horn and C. R. Johnson, *Topics in Matris Analysis* (Cambridge University Press, Cambridge, 1999).
[24] R. Olfati-Saber, J. A. Fax, and R. M. Murray, Proc. IEEE **95**, 215 (2007).
[25] O. E. Rössler, Phys. Lett. A **57**, 397 (1976).
[26] Y. Kuramoto, *Chemical Oscillations, Waves and Turbulence* (Springer-Verlag, Berlin, 1984); S. H. Strogatz, Physica D **143**, 1 (2000).
[27] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[28] P. Erdős and A. Rényi, Publ. Math. (Debrecen) **6**, 290 (1959).
[29] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).
[30] D. Lusseau, K. Schneider, O. J. Boisseau, P. Haase, E. Slooten, and S. M. Dawson, Behav. Ecol. Sociobiol. **54**, 396 (2003).
[31] M. Girvan and M. E. J. Newman, Proc. Natl. Acad. Sci. U.S.A. **99**, 7821 (2002).
[32] W. W. Zachary, J. Anthropol. Res. **33**, 452 (1977).
[33] See http://www.orgnet.com/cases.html.
[34] J. Sun, E. M. Bollt, and T. Nishikawa, Europhys. Lett. **85**, 60011 (2009).
[35] N. Fujiwara and J. Kurths, Eur. Phys. J. B **69**, 45 (2009).