Reconstructing distant interactions of multiple paths between perceptible nodes in dark networks

Xinyu Wang,[1] Yuanyuan Mi,[2] Zhaoyang Zhang,[3] Yang Chen,[4] Gang Hu,[5,*] and Haihong Li[1,†]

[1]*School of Science, Beijing University of Posts and Telecommunications, Beijing 100876, China*

[2]*Center for Neurointelligence, School of Medicine, Chongqing University, Chongqing 400044, China and AI Research Center, Peng Cheng Laboratory, Shenzhen 518005, China*

[3]*Department of Physics, School of Physical Science and Technology, Ningbo University, Ningbo, Zhejiang 315211, China*

[4]*Brainnetome Center and National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China*

[5]*Department of Physics, Beijing Normal University, Beijing 100875, China*

(Received 20 December 2021; revised 30 April 2022; accepted 9 June 2022; published 5 July 2022)

Quantitative research of interdisciplinary fields, including biological and social systems, has attracted great attention in recent years. Complex networks are popular and important tools for the investigations. Explosively increasing data are created by practical networks, from which useful information about dynamic networks can be extracted. From data to network structure, i.e., network reconstruction, is a crucial task. There are many difficulties in fulfilling network reconstruction, including data shortage (existence of hidden nodes) and time delay for signal propagation between adjacent nodes. In this paper a deep network reconstruction method is proposed, which can work in the conditions that even only two nodes (say A and B) are perceptible and all other network nodes are hidden. With a well-designed stochastic driving on node A, this method can reconstruct multiple interaction paths from A to B based on measured data. The distance, effective intensity, and transmission time delay of each path can be inferred accurately.

DOI: 10.1103/PhysRevE.106.014302

I. INTRODUCTION

In recent decades, massive amounts of data in different systems have been continuously generated with the rapid development of various experimental technologies, such as, the huge amounts of data about the neuronal activities in neural system. Based on data accumulation, analytical quantitative computations enter diverse fields and reveal common properties of a wide range of systems in nature and society [1–5]. To describe systems composed by a large number of highly interconnected dynamical units, one approach is to model them as networks whose nodes represent the local dynamical units, and whose links stand for the interactions between these units. Various sets of coupled differential equations are widely used to model the evolution of complex networks [6,7]. The nodes in a complex network receive external inputs and interact with each other through connections, and as a result, the network state evolves to process information [8,9].

The activities of nodes in a practical complex network are often measurable. However, the topological structure of the network is often difficult to detect directly. Thus, exploring the unknown network topology based on available node activities, the so-called network reconstruction [10], has attracted much attention. A number of reconstruction methods have been proposed, including methods based on external perturbation [11,12] and noise-induced correlations [13–18], the adaptive feedback control, phase modeling approach, and robust regression [19–22]. Additionally, various sophisticated statistical inference techniques have entered use for network reconstruction [23].

Network reconstruction still faces lots of challenges. Previous studies mainly focus on solving the challenges caused by the complexity of a network [24,25], the nonlinearity of local dynamics of nodes [22], and the transmission delay between adjacent nodes [26–28]. Recently, the challenge of hidden nodes [29–32] have caused heated discussion. Besides, large reconstruction errors may also arise due to the mutual interactions between nodes via multiple paths and with different distances [24], and the lack of prior information about the network structure [23].

In this paper, we proposed a method for deep network reconstruction. By deep reconstruction we mean inferring the interactions between distant nodes along multiple paths, including the information about the distance, the effective interaction intensities and the transmission time delay in each path. We apply an external driving signal to the network. The driving signal in our method is a well-designed stochastic signal with short correlation time, and it will be described in detail in the following section. Through analyzing the network response to the driving signal, we explore the information about the network structure and the network dynamics. Without knowledge of the underlying system, analyses of data produced by network activity under driving can be of great help to understanding the relations between network structure and network dynamics. It should be emphasized that all the

*ganghu@bnu.edu.cn
†haihongli@bupt.edu.cn

 ©2022 American Physical Society

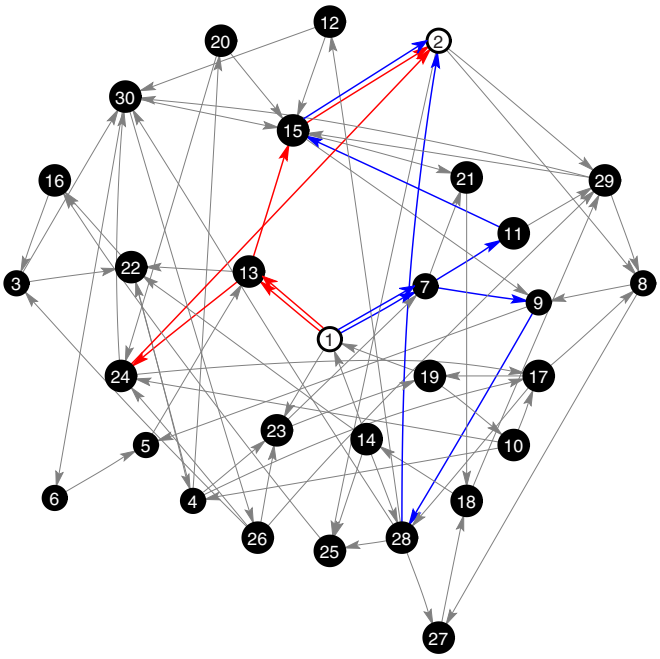

FIG. 1. Schematic diagram of the deep network reconstruction method. The external driving signal is loaded on a node A ($A = 1$ as an example in the following study), and the response activity of an arbitrary node B ($B = 2$) is measurable. All the rest black nodes are hidden nodes. The task is to reconstruct the multiple interaction paths from A to B, such as the shortest path (red lines) and the second shortest paths (blue lines), including information of the corresponding distances, interaction intensities and time delays.

above reconstruction computations can be implemented based on the activities of a limited number of accessible nodes only (even two nodes in our analyses), and the rest are hidden (see Fig. 1).

For the stochastic driving signal, it intuitively plays negative roles in many fields, especially for the data analyses [33]. However, since the end of the last century, it has been shown that a dynamical system that is subject to both periodic forcing and random perturbations may show a resonancelike behavior, called stochastic resonances (SRs) [34]. A SR-type behavior in excitable systems with aperiodic input signals is also presented [35], where it is showed that noise can serve to enhance the response of a nonlinear system to a weak input signal, regardless of whether the signal is periodic or aperiodic. Furthermore, noise can play a positive role in enhancing signals or coherence intensities in nonlinear systems under some critical conditions, called coherence resonance (CR) [36,37].

The paper is organized as follows. In Sec. II, we introduce the deep network reconstruction method theoretically. In Sec. III, we demonstrate the validity of the method and how to numerically perform the algorithm in some simple networks. In Sec. IV, we generalize the method and show explicit results of complex networks with large size and different types of local dynamics. Finally, Sec. V summarizes the main results in our study and gives some insights and perspectives for the potential applications of our method.

II. DEEP NETWORK RECONSTRUCTION METHOD VIA STOCHASTIC DRIVING SIGNAL—THE THEORETICAL ANALYSIS

We consider a networked coupled system of N nodes, which is described by a set of differential equations where all functions $\Phi_i, i = 1, 2, ..., N$ are *a priori* unknown (for convenience, multiple dimensional state variables are not considered in this section):

$$\dot{x}_i(t) = \Phi_i[\vec{X}_i(t)] + \Gamma_i(t), \quad i = 1, 2, ..., N, \tag{1a}$$

where

$$\vec{X}_i(t) = [x_1(t - \tau_{i1}), ..., x_i(t), ..., x_N(t - \tau_{iN})].$$

Φ_i is a general form of the dynamical function of node i, including the settling local dynamics and the interactions from other nodes in the network ($\frac{\partial \Phi_i}{\partial x_j} \equiv 0$ if connection from j is absent). A positive constant τ_{ij} denotes time delay from j to i. A dynamical model which considers the transmission time delay is close to practical systems and is capable of elucidating much richer behaviors in real world. We adopt a Gaussian white noise $\Gamma_i(t)$ for inherent noise disturbances satisfying

$$\langle \Gamma_i(t) \rangle = 0, \langle \Gamma_i(t)\Gamma_i(t + t') \rangle = Q_i \delta(t'), \tag{1b}$$

where Q_i is the noise intensity, a positive constant. The δ function reflects that correlation time of the noise is much shorter than the characteristic time of the system dynamics.

We can measure the variable $x_i(t)$ of node i that changes continuously over time if it is perceptible. Our goal is to reconstruct the interactions between two perceptible nodes based on the measurable data. The other variables, such as the dynamics of all hidden nodes, the connection structure of network, the transmission time delay between adjacent nodes, background noise, etc., are unknown.

For reconstructing interaction from one node (say, node A) to another node (node B), we inject a varying external driving $D_A(t)$ at node A as

$$\dot{x}_A(t) = \Phi_A[\vec{X}_A(t)] + \Gamma_A(t) + D_A(t). \tag{2}$$

Assuming that there is a direct connection from A to B, for node B we have

$$\dot{x}_B(t) = \Phi_B[\vec{X}_B(t)] + \Gamma_B(t), \tag{3}$$

where $\frac{\partial \Phi_B}{\partial x_A} \neq 0$, called B is distance 1 ($d = 1$) from A.

To analyze output response in B, first we take the derivative of two sides of Eq. (3) with respect to t, which gives

$$\begin{aligned}
\ddot{x}_B(t) &= \sum_{i=1}^{N} \frac{\partial \Phi_B[\vec{X}_B(t)]}{\partial x_i(t - \tau_{Bi})} \frac{dx_i(t - \tau_{Bi})}{dt} + \dot{\Gamma}_B(t) \\
&= \sum_{i=1}^{N} \frac{\partial \Phi_B[\vec{X}_B(t)]}{\partial x_i(t - \tau_{Bi})} \Phi_i[\vec{X}_i(t - \tau_{Bi})] \\
&\quad + \sum_{i=1}^{N} \frac{\partial \Phi_B[\vec{X}_B(t)]}{\partial x_i(t - \tau_{Bi})} \Gamma_i(t - \tau_{Bi}) \\
&\quad + \frac{\partial \Phi_B[\vec{X}_B(t)]}{\partial x_A(t - \tau_{BA})} D_A(t - \tau_{BA}) + \dot{\Gamma}_B(t),
\end{aligned} \tag{4}$$

where $\ddot{x}_B$ is the second derivative in time of x_B and $\dot{\Gamma}_B$ denotes an irrelevant noise-induced term with no practical meaning. It can be concluded that there is only one driving-induced term which builds a direct quantitative relationship with coupling function, provided $D_A(t)$ is independent of the system,

$$\frac{\partial \Phi_B[\vec{X}_B(t)]}{\partial x_A(t-\tau_{BA})} D_A(t-\tau_{BA}). \tag{5}$$

We collect the response data of node B with a fixed time interval ($\Delta t \ll 1$), denoted as

$$[x_B(t_1), x_B(t_2), ..., x_B(t_l), ..., x_B(t_L)], \tag{6}$$

where $t_{l+1} - t_l = \Delta t$, $T = L\Delta t$. From data sequence, the first and the second derivative of x_B are given by

$$\dot{x}_B(t_l) \approx x_B^{(1)}(t_l) = \frac{x_B(t_{l+1}) - x_B(t_l)}{\Delta t},$$

$$\ddot{x}_B(t_l) \approx x_B^{(2)}(t_l) = \frac{x_B^{(1)}(t_{l+1}) - x_B^{(1)}(t_l)}{\Delta t},$$

and for $\mu \geqslant 1$

$$\frac{d}{dt^{\mu+1}} x_B(t_l) \approx x_B^{(\mu+1)}(t_l) = \frac{x_B^{(\mu)}(t_{l+1}) - x_B^{(\mu)}(t_l)}{\Delta t}. \tag{7}$$

To detect the driving component D_A, we calculate the correlation function between the driving signal and $x_B^{(2)}$ where $t' \in [0, T]$

$$\begin{aligned} E_{AB}^{1,2}(t') &= \frac{1}{T-t'} \int_0^{T-t'} D_A(t) x_B^{(2)}(t+t') dt \\ &= \frac{1}{T-t'} \int_0^{T-t'} \frac{\partial \Phi_B[\vec{X}_B(t)]}{\partial x_A(t-\tau_{BA})} D_A(t) D_A(t+t'-\tau_{BA}) dt \\ &\quad + E_{\Gamma_B}^1 + E_{\Phi_B}^1. \end{aligned} \tag{8}$$

This yeilds an intensity spectrum for varying t', which is directly related to the autocorrelation property of driving $D_A(t)$. There are also two terms $E_{\Gamma_B}^1$ and $E_{\Phi_B}^1$ which are evidently unrelated to t'.

For the term of $E_{\Gamma_B}^1$, the emergence of noise is bound to cause a random deviation. From Eqs. (3) and (4), it is known that the leading deviation component is given by

$$E_{\Gamma_B}^1 = \left\langle D_A \Gamma_B^{(1)} \right\rangle = \left\langle D_A(t) \frac{\Gamma_B(t+\Delta t) - \Gamma_B(t)}{\Delta t} \right\rangle. \tag{9a}$$

Additionally, for the term of $E_{\Phi_B}^1$, the system dynamics can induce a considerable bias,

$$E_{\Phi_B}^1 = \left\langle D_A \sum_{i=1}^{N} \left[\frac{\partial \Phi_B(\vec{X}_B)}{\partial x_i} \Phi_i(\vec{X}_i) \right] \right\rangle. \tag{9b}$$

According to Eq. (8), the correlation function $E_{AB}^{1,2}(t')$ can be calculated based on the measured data. It is expected that the driving signal can be well designed so that both terms, $E_{\Gamma_B}^1$ and $E_{\Phi_B}^1$, approach to infinitesimal quantities. The driving signal having the following characteristics can meet the above requirement.

(a) The driving signal is sampled from a series of quantified *stochastic* numbers $D_A(t_1), D_A(t_2)..., D_A(t_l), ..., D_A(t_L)$, i.e.,

$$D_A(t) = D_A(t_l), t \in (t_{l-1}, t_l],$$

where the tunable time interval $t_l - t_{l-1}$ should be much smaller than the system characteristic time (so $D_A(t)$ is called *fast varying* driving).

(b) The statistical *mean* of the driving signal is *zero*, i.e.,

$$\langle D_A(t) \rangle = \lim_{t_L \to \infty} \frac{1}{t_L} \sum_{l=1}^{L} D_A(t_l)(t_l - t_{l-1}) \to 0.$$

The zero-mean property can help to weaken the effect of the driving signals on the system and reduce the bias in the correlation function. E_{Φ_B} is proved to be equivalently infinitesimal as $(t_l - t_{l-1})$ by theoretical derivation and can be ignored,

$$E_{\Phi_B} \to O(t_l - t_{l-1}).$$

More detailed analysis are shown in Appendix A. In addition, the dynamics of the systems not only depend on the current state but also on the previous states, which may induce complicated or practically important behaviors of system dynamics (e.g., excitability or chaotic dynamics). Therefore, the reconstruction of time delay τ_{BA} is essential. According to Eq. (8), to distinguish time delay in the autocorrelation the driving signal should have the following characteristics,

(c) The driving signal is *nonperiodic*.

(d) The driving signal has *short correlation time*, i.e.,

$$\begin{aligned} \langle D_A(t_m) D_A(t_n) \rangle &\neq 0, m = n, \\ \langle D_A(t_m) D_A(t_n) \rangle &= 0, m \neq n. \end{aligned}$$

It is interesting that white noise, which is common and inherent in many practical systems, satisfies all the above requirements, such as the neural circuits and technology networks. Many previous studies on network reconstruction show that the network can generate rich distinctive data in the presence of white noises. Thus, we designed a stochastic sequence to simulate noiselike driving for network reconstruction:

$$D_A(t) = D_A(t_l), t \in (l\Delta t_D - \Delta t_D, l\Delta t_D), \tag{10a}$$

where the signal changes with a fixed time interval Δt_D. The value of $D_A(t_l)$ is randomly sampling from a Gaussian distribution

$$D_A(t_l) \sim N(0, Q_{DA}/\Delta t_D), \tag{10b}$$

where Q_{DA} is the intensity. For simplification, we set $\Delta t_D = \Delta t$ in our study, and hence,

$$\lim_{L \to \infty} \frac{1}{L} \sum_{l=1}^{L} D_A(t_l) \to 0,$$

$$\lim_{L \to \infty} \frac{1}{L} \sum_{l=1}^{L} D_A(t_l) D_A(t_l) \Delta t \to Q_{DA},$$

and for any $k > 0$,

$$\lim_{L \to \infty} \frac{1}{L} \sum_{l=1}^{L-k} D_A(t_l) D_A(t_l + k\Delta t) \to 0.$$

Therefore, based on the well-designed stochastic driving signal, Eq. (8) can be rewritten as

$$E_{AB}^{1,2}(t') = \begin{cases} O(\Delta t) & t' \neq \tau_{BA}, \\ I_{BA}(A \to B)Q_{DA} + O(\Delta t) & t' = \tau_{BA}, \end{cases} \tag{11}$$

where

$$I(A \to B) = \left\langle \frac{\partial \Phi_B[\vec{\boldsymbol{X}}_B(t)]}{\partial x_A(t - \tau_{BA})} \right\rangle, \tag{12}$$

where $I(A \to B)$ is the effective intensity of the interaction from A to B. We can reconstruct the fundamental information of this adjacent link, including the accurate connection strength and the time delay. However, the interaction between two direct-connected nodes can only provide limited information compared to the overall system, while it has always been a challenging problem to reconstruct indirect connections in the network. Intuitively, an indirect interaction from A to B cannot be inferred with only the measurable data of the destination node B, since the dynamics of node is also strongly influenced by other hidden nodes.

Considering a path $P : A \to j_1 \to B$ with one intermediate node j_1, i.e., the distance from A to B is $d = 2$, and

$$\frac{\partial \Phi_i(\vec{\boldsymbol{X}}_i(t))}{\partial x_j(t - \tau_{ij})} \begin{cases} \neq 0 & (i, j) = (B, j_1) \text{ or } (j_1, A), \\ = 0 & \text{others.} \end{cases}$$

In this case the driving-induced term disappears in $\ddot{x}_B$ due to $\partial \Phi_B[\vec{\boldsymbol{X}}_B(t)]/\partial x_A(t - \tau_{BA}) \equiv 0$. So we calculate the higher-order derivatives of $x_B(t)$, which is given by,

$$\begin{aligned} x_B^{(3)}(t) \approx [\ddot{x}_B(t)]^{(1)} = \frac{d}{dt}\left(\frac{\partial \Phi_B[\vec{\boldsymbol{X}}_B(t)]}{\partial x_{j_1}(t - \tau_{Bj_1})} \right) \Phi_{j_1}[\vec{\boldsymbol{X}}_{j_1}(t - \tau_{Bj_1})] + \frac{\partial \Phi_B[\vec{\boldsymbol{X}}_{j_1}(t)]}{\partial x_{j_1}(t - \tau_{Bj_1})} \frac{\partial \Phi_{j_1}[\vec{\boldsymbol{X}}_{j_1}(t - \tau_{Bj_1})]}{\partial x_A(t - \tau_{Bj_1} - \tau_{j_1A})} \{ \Phi_A[\vec{\boldsymbol{X}}_A(t - \tau_{Bj_1} - \tau_{j_1A})] \\ + \Gamma_A(t - \tau_{Bj_1} - \tau_{j_1A}) + D_A(t - \tau_{Bj_1} - \tau_{j_1A}) \}. \end{aligned} \tag{13}$$

And the driving signal D_A can propagate along the path $A \to j_1 \to B$ via the following driving-induced term,

$$\frac{\partial \Phi_B[\vec{\boldsymbol{X}}_{j_1}(t)]}{\partial x_{j_1}(t - \tau_{Bj_1})} \frac{\partial \Phi_{j_1}[\vec{\boldsymbol{X}}_{j_1}(t - \tau_{Bj_1})]}{\partial x_A(t - \tau_{Bj_1} - \tau_{j_1A})} D_A(t - \tau_{Bj_1} - \tau_{j_1A}).$$

To detect the driving force in $x_B(t)$, we calculate the correlation between $D_A(t)$ and $x_B^{(3)}(t)$, which is given by

$$\begin{aligned} E_{AB}^{1,3}(t') &= \frac{1}{T - t'} \int_0^{T - t'} D_A(t) x_B^{(3)}(t + t') dt \\ &= \begin{cases} O(\Delta t) & t' \neq \tau_{Bj_1} + \tau_{j_1A}, \\ I(A \to j_1 \to B)Q_{DA} + O(\Delta t) & t' = \tau_{Bj_1} + \tau_{j_1A}, \end{cases} \end{aligned} \tag{14}$$

where

$$I(A \to j_1 \to B) = \left\langle \frac{\partial \Phi_B[\vec{\boldsymbol{X}}_{j_1}(t)]}{\partial x_{j_1}(t - \tau_{Bj_1})} \frac{\partial \Phi_{j_1}[\vec{\boldsymbol{X}}_{j_1}(t - \tau_{Bj_1})]}{\partial x_A(t - \tau_{Bj_1} - \tau_{j_1A})} \right\rangle.$$

Finally, we consider the case $d > 2$, where there are $d - 1$ different hidden nodes along the path from A to B, where number ν denotes the distance of intermediate node j_ν away from A

$$P : A \to j_1 \to \ldots j_\nu \to \ldots \to j_{d-1} \to B. \tag{15a}$$

The structural effective interaction intensity from A to B along a given path P is defined

$$I_{BA}(P) = \left\langle \frac{\partial \Phi_B}{\partial x_{j_{d-1}}} \cdot \prod_{\nu=2}^{d-1} \frac{\partial \Phi_{j_\nu}}{\partial x_{j_{\nu-1}}} \cdot \frac{\Phi_{j_1}}{\partial x_A} \right\rangle, \tag{15b}$$

where for $i, j = j_1, j_2, .., j_{d-1}, B$

$$\frac{\partial \Phi_i}{\partial x_j} = \frac{\partial \Phi_i[\vec{\boldsymbol{X}}_i(t + \tau_{iA})]}{\partial x_j[t + \tau_{jA}(P)]},$$

and time delay for j_ν through P is

$$\tau_{j_\nu A}(P) = \tau_{j_1 A} + \sum_{\mu=2}^{\nu} \tau_{j_\mu j_{\mu-1}}. \tag{15c}$$

To detect the driving force component, we calculate the correlation between the driving signal $D_A(t)$ and the $d + 1$th derivative of $x_B(t)$, which is given by

$$E_{AB}^{1,d+1}(t') \approx \begin{cases} O(\Delta t) & t' \neq \tau_{BA}(P), \\ I_{BA}(P)Q_{DA} + O(\Delta t) & t' = \tau_{BA}(P). \end{cases} \tag{16}$$

The calculation of $E_{AB}^{1,d+1}(t')$ is only dependent on the data of $x_B(t)$ and $D_A(t)$. According to the theoretical deduction, there is a discontinuity at $t' = \tau_{BA}$, which is called the characteristic discontinuity of the certain path P.

Note that there are two more terms $E_{\Gamma_B}^d$ and $E_{\Phi_B}^d$ in the correlation function. The leading terms of the random deviation caused by noises and the bias caused by the system dynamics are given by

$$E_{\Gamma_B}^d = \langle D_A(t) \Gamma_B^{(d)} \rangle, \tag{17a}$$

$$E_{\Phi_B}^d = \left\langle D_A(t) \sum_{i,j=1}^{N} F_\Phi^d \right\rangle, \tag{17b}$$

where F_Φ^d is a simplified function given as

$$F_\Phi^d \left[\Phi_i, \frac{\partial \Phi_i}{\partial x_j}, \frac{d}{dt}\left(\frac{\partial \Phi_i}{\partial x_j} \right), ..., \frac{d^d}{dt^d}\left(\frac{\partial \Phi_i}{\partial x_j} \right) \right].$$

Under the condition of the well-designed stochastic driving signal, these two irrelevant quantities become small and negligible for long time averages (more details are shown in Appendix A).

Due to the discretization of measurement time, the above correlation can be rewritten as

$$E_{AB}^{1,d+1}(k\Delta t) = \frac{\Delta t}{T - k\Delta t}\sum_{l=1}^{L-k} D_A(t_l)x_B^{(d+1)}(t_l + k\Delta t), \quad (18a)$$

where Δt is the time interval in the recording process. The reconstructed interaction intensity can be represented by

$$\tilde{I}_{BA}(P) = E_{AB}^{1,d+1}[\tilde{\tau}_{BA}(P)]/Q_{DA}, \quad (18b)$$

where $\tilde{\tau}_{BA}(P)$ is the experimental time delay [ceil(x) means rounding up number x]

$$\tilde{\tau}_{BA}(P) = \Delta t \cdot \text{ceil}[\tau_{BA}(P)/\Delta t]. \quad (18c)$$

The systematical deep network reconstruction proposed above can be summarized as follows. To infer the information of an arbitrary path from A to B, we load a signal $D_A(t)$ on node A, and record the response data of B $[x_B(t)]$. We then calculate the correlation function $E_{AB}^{1,n}(t')$ $(t' \in [0, T])$. If a singularity appears at $t' = \tilde{\tau}_{BA}(P)$, it means that the distance is $d = n - 1$ and there are $n - 2$ hidden nodes in the path from A to B. The transmission delay $\tilde{\tau}_{BA}(P)$, and the effective intensity $\tilde{I}_{BA}(P)$ can be calculated according to Eq. (18). Furthermore, multiple paths can be detected if more singularities are discovered. A schematic figure of a complex network is shown in Fig. 1, in which the multiple paths to be reconstructed are colored. In the next section, we apply the deep network reconstruction to some simple networks and demonstrate the validity of our method with numerical simulations.

III. NUMERICAL DEMONSTRATION OF DEEP NETWORK RECONSTRUCTION METHOD USING SIMPLE NETWORKS

In this section, we demonstrate the feasibility of the deep network reconstruction method in networks with small sizes and linear dynamics.

A. Unnoisy linear system

We consider a linear system without the inherent noise where an external driving signal is loaded on node A ($A = 1$ as an example)

$$\dot{x}_{A=1}(t) = -\alpha x_1(t) + \sum_{j\neq 1} M_{1j}x_j(t - \tau_{1j}) + D_1(t), \quad (19a)$$

and for the other nodes $i \neq A$

$$\dot{x}_i(t) = -\alpha x_i(t) + \sum_{j\neq i} M_{ij}x_j(t - \tau_{ij}), \quad (19b)$$

where $\alpha = 4$ and $\hat{\boldsymbol{M}}$ is the connection matrix. Node B ($B = 2$) is perceptible and the interaction paths from A to B are tested for deep reconstruction. Both the values of measurement intervals (Δt) and the loading interval of driving signal (Δt_D) are set to be the same, $\Delta t_D = \Delta t = 10^{-2}$.

First, we consider a network named Net1. The connection matrix is set to be

$$\text{Net1} : \hat{\boldsymbol{M}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1.2 & 1.5 \\ 1.2 & 0 & 0 & 0 \\ 1.5 & 0 & 0 & 0 \end{pmatrix}.$$

The transmission delay matrix is shown by

$$\hat{\boldsymbol{\tau}} = \begin{pmatrix} - & - & - & - \\ - & - & 0.835 & 0.779 \\ 0.739 & - & - & - \\ 0.604 & - & - & - \end{pmatrix}.$$

In this network, there are two different paths from node $A = 1$ to $B = 2$ $P_1 : 1 \to 3 \to 2$, $P_2 : 1 \to 4 \to 2$ as shown in Fig. 2(a). From Eqs. (15b) and (15c), the theoretical values of the effective intensities and time delays in these paths are given by

$$I_{BA}(P_1) = M_{31}M_{23} = 1.44, \quad I_{BA}(P_2) = M_{41}M_{24} = 2.25,$$
$$\tau_{BA}(P_1) = \tau_{31} + \tau_{23} = 1.574, \quad \tau_{BA}(P_2) = \tau_{41} + \tau_{24} = 1.383.$$

The network activity fluctuates around $x_i^* = 0, i = 1, 2, ..., N$, which is the stable state of the network. The source variable $x_1(t)$ fluctuates in a wider range because of the driving effect while others show weak responses, as shown in Fig. 2(b). We calculate the correlation $E_{AB}^{1,n}(t')$ at different n and reasonable range of t', two characteristic singularities emerge until $n = 3$ as shown in Fig. 2(c), which means there are two different paths from A to B, and the distances are both $d = 2$. The positions t' of these singularities correspond to the transmission time delay in these paths, $\tilde{\tau}_{BA}(P_1) = 1.57$, $\tilde{\tau}_{BA}(P_2) = 1.38$. The corresponding effective intensity of the path can be calculated via the average of differences between the amplitude of singularity and the background level, i.e.,

$$\tilde{I}_{BA}(P_1) = \frac{1.3031 - 0 + 1.3031 - (-0.1827)}{2} = 1.3944,$$
$$\tilde{I}_{BA}(P_2) = \frac{2.1085 - 0 + 2.1085 - (-0.2396)}{2} = 2.2283.$$

The numerical reconstruction results are summarized in Table I, where the accuracy of the intensity is defined as ratio $= \tilde{I}_{BA}/I_{BA}$, very close to 1.

According to the above numerical results, a sketch map of the reconstructed network is shown in Fig. 2(d). Compared with the structure Fig. 2(a), two paths are well reconstructed, including two different sandwiched hidden nodes.

In Net1, we reconstruct multiple paths with the same distance. In the following part, we demonstrate that our method works in the condition of multiple paths with different distances. We make some changes to the structure of Net1 and generate a network with $N = 5$ named as Net2,

$$\text{Net2} : \hat{\boldsymbol{M}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1.2 & 0 & 1.5 \\ 1.2 & 0 & 0 & 0 & 0 \\ 1.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.5 & 0 \end{pmatrix}.$$

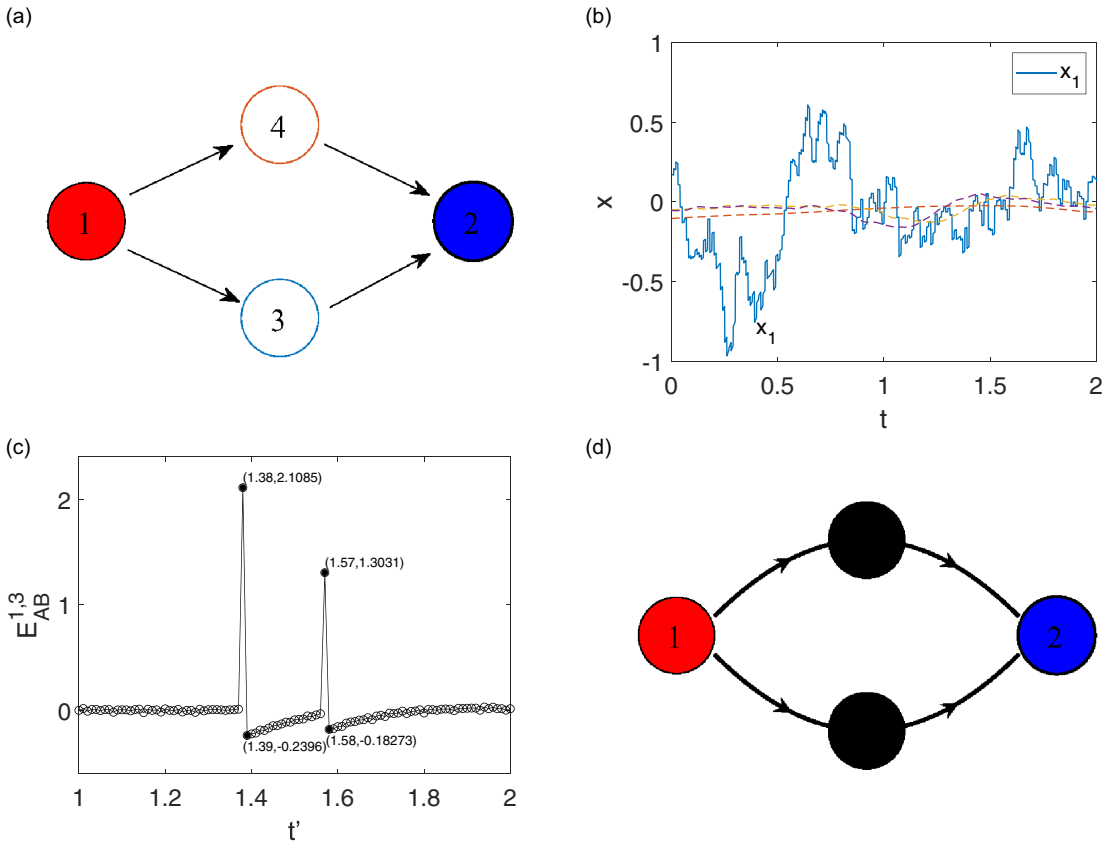

FIG. 2. Reconstruction results of network Net1. The driving intensity is $Q_{DA} = 1$. The measurement time length is $T = 10\,000$, with $\Delta t = 0.01$. (a) Schematic diagram of Net1. (b) The response sequences of four nodes in Net1, where the activity of $x_1(t)$ fluctuates in a wide range due to the driving signal, while the activities of all other nodes stay around the stable state. (c) Plot of the correlation function $E_{AB}^{1,3}(t')$ in the range $t' \in [1, 2]$. The two characteristic singularities of at $t' = 1.38$ and $t' = 1.57$ mean that there are two different paths from 1 to 2, whose amplitude value are related to the effective intensity. And the abscissa position of t' suggests the transmission delay of the corresponding path, respectively. (d) Sketch map of the reconstructed network.

The parameters of time delay are as follows:

$$\hat{\boldsymbol{\tau}} = \begin{pmatrix} - & - & - & - & - \\ - & - & 0.835 & - & 0.720 \\ 0.739 & - & - & - & - \\ 0.604 & - & - & - & - \\ - & - & - & 0.945 & - \end{pmatrix}.$$

In this case an extra hidden node 5 is considered. There are two paths from A ($A = 1$) to B ($B = 2$) with different distances, i.e., $P_1 : 1 \to 3 \to 2$ and $P_2 : 1 \to 4 \to 5 \to 2$, as shown in Fig. 3(a). To reconstruct the multiple paths, the correlations $E_{AB}^{1,n}(t')$ are calculated based on data from Net2. We find that a singularity emerges in $E_{AB}^{1,3}(t')$ at $t' = 1.57$ first, as shown in Fig. 3(b), which means there is a shortest path with $d = 2$, and a new singularity emerges in $E_{AB}^{1,4}(t')$ at $t' = 2.27$ as shown in Fig. 3(c), informing there is a second shortest path with $d = 3$. For convenience, we color the two characteristic singular peaks in Figs. 3(b) and 3(c) to distinguish the shortest path and the second shortest path, for the shortest path of $d = 2$ and for the second shortest path of $d = 3$. The hollow points represent the position of theoretical value I_{BA} and τ_{BA}. Note that the value of $E_{AB}^{1,4}(t')$ at $t' = 1.57$ is much larger than that in $E_{AB}^{1,3}(t')$. This ia because correlation function comes from the same data of $x_B(t)$, and correlation in $E_{AB}^{1,3}(t')$ is enlarged by $1/\Delta t$ in $E_{AB}^{1,4}(t')$, not considered as a new singularity. Similarly, the same driving signal exist in multiple paths with different distances and time delays, which can cause random background perturbation in $E_{AB}^{1,4}(t')$. Thus, we only consider the amplitude of singularity at $t' = 2.27$ as the effective intensity of the second shortest path because of the background perturbation. The reconstructed results are shown in Table II, and the accuracy of intensities for both paths is very close to 1. According to the above numerical simulation, a sketch map of the reconstructed network is shown in Fig. 3(d).

TABLE I. Reconstructed results of Net1.

	d	I_{BA}	$\tilde{I}_{BA}$	ratio	τ_{BA}	$\tilde{\tau}_{BA}$
P_1	2	1.44	1.3944	0.9683	1.574	1.57
P_2	2	2.25	2.2283	0.9904	1.383	1.38

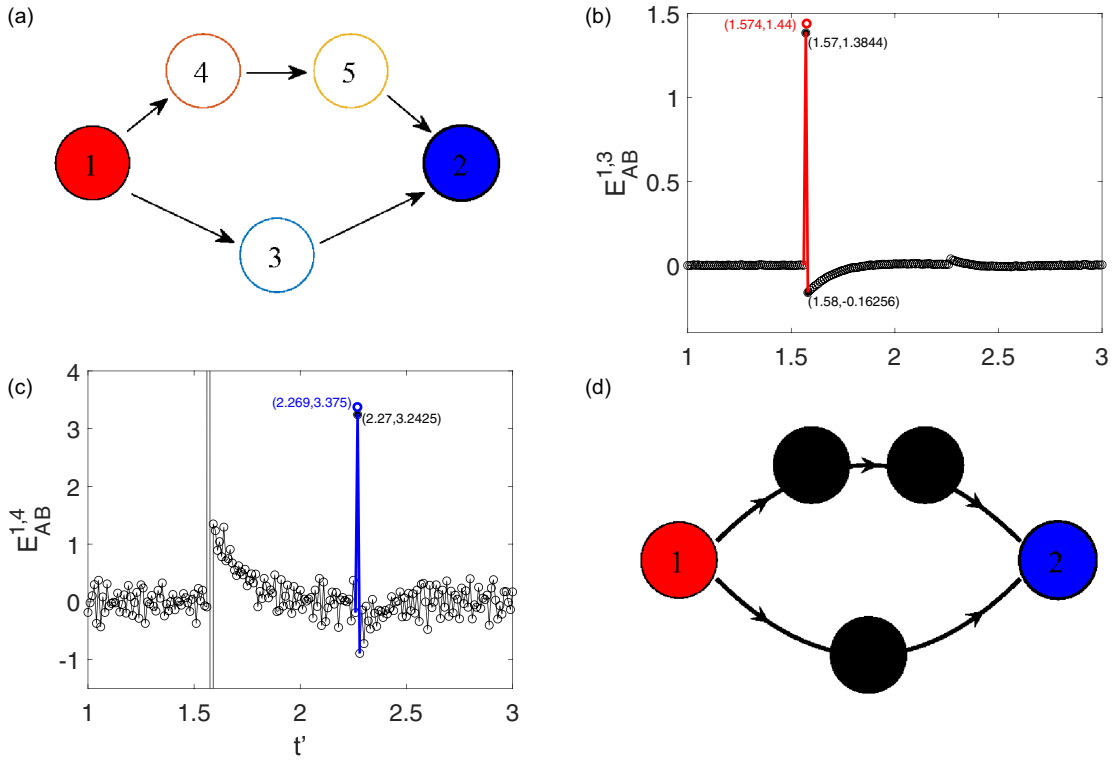

FIG. 3. Reconstruction results of Net2. The same as Fig. 2, $Q_{DA} = 1$, $T = 10\,000$ and $\Delta t = 0.01$. (a) Schematic diagram of Net2. (b, c) The correlations $E^{1,3}_{AB}(t')$ and $E^{1,4}_{AB}(t')$ in the range of $t' \in [1, 3]$. In panel (b) a characteristic singularity (colored in red) emerges at $t' = 1.57$ first, which means the inferred distance of the shortest path is $d_s = 2$, and the amplitude is related to the inferred effective intensity. The circle indicates the actual interaction intensity $I_{BA}(P_1)$ and time delay $\tau_{BA}(P_1)$. $I_{BA}(P_1) = 1.44$ and $\tau_{BA}(P_1) = 1.574$ are represented by the abscissa and ordinate of the circle. (c) Another characteristic singularity (colored in blue) emerges at $t' = 2.27$ in $E^{1,4}_{AB}(t')$, which means the inferred distance of the second shortest path is 3, and the amplitude is related to the inferred effective intensity. The circle indicates the actual interaction intensity $I_{BA}(P_2)$ and time delay $\tau_{BA}(P_2)$. For the shortest and the second shortest path, theoretical and experimental values meet well with each other.

Finally, we remove node 3 of Net2 and obtain Net3:

$$\text{Net3}: \hat{\boldsymbol{M}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1.2 & 0 & 0 & 0 & 1.5 \\ 0 & 0 & 0 & 0 & 0 \\ 1.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.5 & 0 \end{pmatrix}.$$

The two different paths are changed as shown in Fig. 4(a), i.e., $P_1 : 1 \to 2$ and $P_2 : 1 \to 4 \to 5 \to 2$, where the difference between two path distances grow larger $d_1 = 1$ and $d_2 = 3$. The shortest path P_1 can be reconstructed via $E^{1,2}_{AB}$ as shown in Fig. 4(b). However, the components of stochastic disturbance in Fig. 4(d) are enlarged once more [much larger than in Fig. 3(c)]. The higher order correlations are severely disturbed so that $E^{1,4}_{AB}$ are very noisy and no singularity can be found as shown in Figs. 4(c) and 4(d) in Net3. In the following part of this paper we focus on inference of the shortest and the second shortest paths with distance difference not more than 1 (for example, if the distance of the shortest path is d_s, then the distance of the second shortest path is $d_s + 1$). In principal, the method is applicable for reconstructing the second shortest paths of farther distances (larger than $d_s + 1$) with often much more measured data.

TABLE II. Reconstructed results of Net2.

	d	I_{BA}	$\tilde{I}_{BA}$	ratio	τ_{BA}	$\tilde{\tau}_{BA}$
P_1	2	1.44	1.4657	1.0178	1.574	1.57
P_2	3	3.375	3.2425	0.9607	2.269	2.27

B. Linear system with noises

A linear system with noise is considered, which is given by

$$\dot{x}_1(t) = -4x_1(t) + \sum_{j\neq 1} M_{1j}x_j(t-\tau_{1j}) + \Gamma_1(t) + D_1(t),$$

$$\dot{x}_i(t) = -4x_i(t) + \sum_{j\neq i} M_{ij}x_j(t-\tau_{ij}) + \Gamma_i(t),\ i \neq 1, \quad (20)$$

where the external signal is loaded on node $A = 1$, the connection matrix is the same as Net2, and Γ_i represents the internal noise on node i.

The shortest interaction path $P_1 : 1 \to 3 \to 2$ with $d = 2$ can be reconstructed as shown in Fig. 5(a). However, the other singularity at $t' = 2.27$ [similar to Fig. 3(c)] related to the second shortest path $P_2 : 1 \to 4 \to 5 \to 2$ disappears due to the noise effect as shown in Fig. 5(b). As the theoretical analysis, reconstruction errors can arise from the noise effect

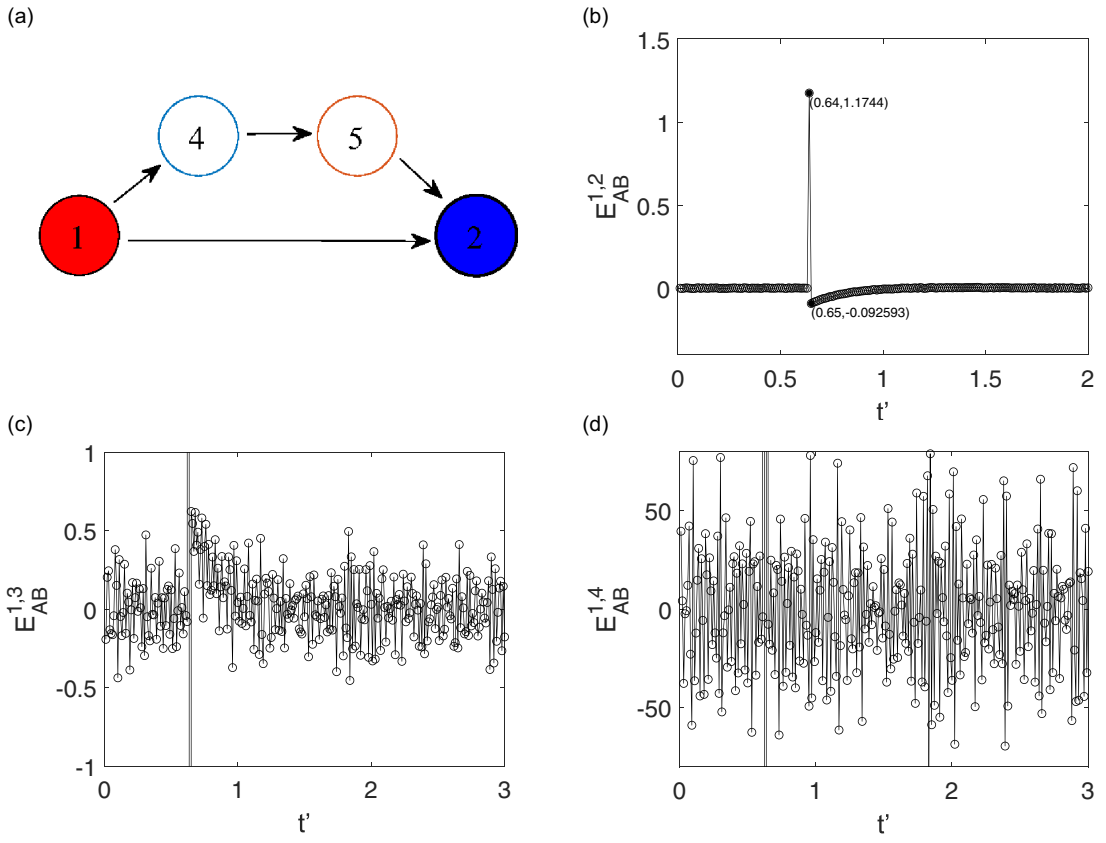

FIG. 4. Reconstruction results of Net3. $Q_{DA} = 1$, $T = 10\,000$ and $\Delta t = 0.01$. (a) Schematic diagram of Net3 (b–d) Plot of the correlation functions $E^{1,2}_{AB}(t')$, $E^{1,3}_{AB}(t')$, $E^{1,4}_{AB}(t')$ in the range of $t' \in [1, 3]$. In panel (b) there is a characteristic singularity at $t' = 0.64$ in $E^{1,2}_{AB}(t')$, which means the distance of the shortest path is $d_s = 1$. In panels (c) and (d) no characteristic singularity can be detected. The singularity related to the second shortest path is embedded in the noisy background in $E^{1,3}_{AB}(t')$, $E^{1,4}_{AB}(t')$. The reconstruction of the $d = 3$ paths fails.

[Eq. (18a)]. There are some small fluctuations in $E^{1,3}_{AB}(t')$ as shown Fig. 5(a), and the fluctuations are largely enhanced in $E^{1,4}_{AB}(t')$ to be comparable with the corresponding singularity value, making reconstruction of the second shortest path fail.

Theoretically, our method is applicable for eliminating the effects of intrinsic noises under the condition of sufficient available data. However, the amounts of required data grow exponentially as the distance of path increases, which causes difficulties for data collection, storage and calculation in practice. Fortunately, from Eq. (17a) it is evident that one can reduce noise-induced deviation effectively by simply increasing Δt. So we can reduce the noise-induced deviation from this direction. The original dataset can be simply rearranged by replacing Δt to $\Delta t' = k\Delta t$ (no additional data is required)

$$x'_i(t_1) = x_i(t_k), x'_i(t_2) = x_i(t_{2k}), \ldots, x'_i(t_{L/k}) = x_i(t_L),$$

$$D'(t_1) = \sum_{l=1}^{k} D(t_l), D'(t_2) = \sum_{l=k+1}^{2k} D(t_l),$$

$$\ldots, \quad D'(t_{L/k}) = \sum_{l=L-k+1}^{L} D(t_l). \tag{21}$$

To weaken the noise effect in our study, we increase $\Delta t = 0.01$ to be $\Delta t = 0.02$ and $\Delta t = 0.04$, based on the same dataset. Then we recalculate $E^{1,4}_{AB}(t')$ with rearranged data of x'_B and D'_A, shown in Fig. 5(c) ($\Delta t = 0.02$) and in Fig. 5(d) ($\Delta t = 0.04$). The fluctuations are greatly weakened. The singularity related to the second shortest path in $E^{1,4}_{AB}(t')$ emerges as shown in Fig. 5(d). The difference is that due to the data rearrangement there are more than one points together contribute to the mutation. We properly sum over all these singular values for correct reconstruction and more details are shown in Appendix B. The reconstructed results of Net2 in the presence of noises are shown in Table III, which is satisfactory.

All the above reconstruction results verify the theoretical predictions of the deep reconstruction algorithms. We demonstrate the effectiveness of our method in the linear system, and show that distance, effective intensity, and time delay of multiple paths can be well inferred in both unnoisy and noisy conditions, with the help of well designed stochastic driving signal.

IV. INFERRING INTERACTIONS OF COMPLEX NETWORKS, SIMULATIONS AND RESULTS

In the above section, linear networks with simple structure are considered for reconstruction, where specific numerical calculations are emphasized. In this section, we will demonstrate the validity of deep network reconstruction in large networks with various nonlinear local dynamics of nodes,

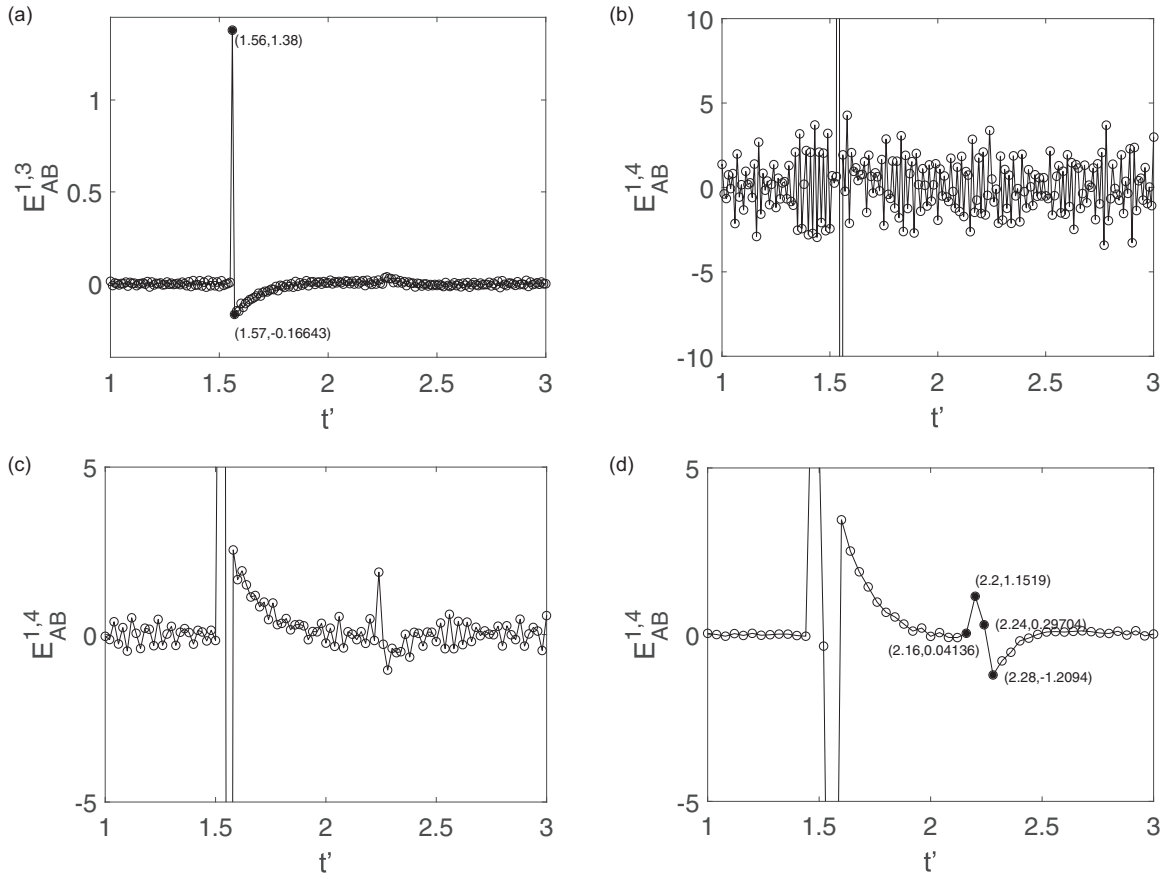

FIG. 5. Reconstruction results of Net2 with the presence of inherent noises ($Q_i = 10^{-6}$, for $i = 1, 2, 3, 4, 5$). The driving intensity is $Q_{DA} = 1$. The measurement time length is $T = 100\,000$. (a) We set $\Delta t = 0.01$ in panel (a). Similar to Fig. 3(b), there is a characteristic singularity at $t' = 1.56$ related to the shortest path. (b–d) The correlation $E^{1,4}_{AB}(t')$ with $\Delta t = 0.01$ in panel (b), $\Delta t = 0.02$ in panel (c) and $\Delta t = 0.04$ in panel (d), respectively. The singularity is completely embedded in the noisy disturbances in panel (b), and appear vaguely in panel (c) while within large fluctuating background. In panel (d) there are several singular points colored in black. These points together form a mutation peak, which is related to the second shortest path, i.e., $P : 1 \to 4 \to 5 \to 2$. This multi-node singularity comes from data rearrangement Eq. (21). The second shortest path can be correctly reconstructed as shown in Table III through summation of the multiple singularities (detailed analysis are shown in Appendix B).

including chaotic and oscillatory dynamics. We will show the reconstruction results in more complicated network structure where nodes are randomly linked with heterogeneous intensities and time delays. The external driving signal $D_A(t)$ can be loaded on an arbitrary perceptible node A, set $A = 1$. Our goal is to reconstruct the distance, effective intensity and time delay of interaction paths from A to an arbitrary node $B(B \neq A)$, under the condition that only nodes A and B are perceptible.

The reconstruction for an arbitrary path from A to B is carried out as follows. First, data of the driving signal $D_A(t)$ and node $x_B(t)$ are recorded simultaneously with a fixed time interval Δt for a duration T. Second, based on the collected data, correlation functions $E^{1,\nu}_{AB}$ are calculated in suitable range of $t' \in [0, T]$ and increasing $\nu = 2, 3, 4, ...$ successively. When a characteristic discontinuity appears first at certain $\nu = \mu_1$ and $t' = \tau_1$, the shortest distance $d_s(A \to B)$ is determined $d_s = \mu_1 - 1$. Further increasing $\nu = \mu_1 + 1, \mu_1 + 2, ...$ until a new discontinuity appears in $E^{1,\nu}_{AB}$ with another $\nu = \mu_2$, the second shortest distance is determined as $d = \mu_2 - 1$. As is discussed in the previous section, driving-induced fluctuations in correlation functions can be greatly amplified while detecting the second shortest distances especially in cases $\mu_2 \geqslant \mu_1 + 2$. In this section we only perform the reconstruction of the paths of the shortest distances with d_s and the second shortest distances with $d_s + 1$, since the reconstruction of longer distances needs too large amount of data. With correctly inferred distances, at last we can search all singularities, each of which infers the time delay $\tilde{\tau}_{BA}$ and effective intensity $\tilde{I}_{BA}$ of one specific path from A to B. In this section, if not specified, time step of the measurement is set $\Delta t = 0.002$.

TABLE III. Reconstructed results of Net2 in the presence of noises.

	d	I_{BA}	$\tilde{I}_{BA}$	ratio	τ_{BA}	$\tilde{\tau}_{BA}$
P_1	2	1.44	1.44335	1.0023	1.574	1.56
P_2	3	3.375	3.07154	0.91008	2.269	2.24

In the following we consider three different networks with distinctive local node dynamics: excitable Bär network, chaotic Lorenz network and gene regulatory network.

A. The excitable complex networks

A considerable amount of works in the field of biological, chemical, and physical systems have been dedicated to excitable media. The interest in this topic is motivated by the fact that wave propagation in these media provides an efficient mechanism for communication between distant locations. Seminal examples include the conduction of electrical impulses along nerve axons. A modified version of FitzHugh-Nagumo model, called the Bär Model is investigated in this paper [38], described as follows ($\delta_{1,i\neq 1}=0$, $\delta_{1,1}=1$):

$$\frac{du_i}{dt} = -\frac{1}{e}u_i(u_i-1)\left(u_i-\frac{v_i+b}{a}\right)+W_i+D_1\delta_{1,i}, \quad (22a)$$

$$\frac{dv_i}{dt} = f(u_i)-v_i, \quad (22b)$$

$$f(u_i)=\begin{cases}0 & u_i<\frac{1}{3},\\ 1-6.75u_i(u_i-1)^2 & \frac{1}{3}\leqslant u_i\leqslant 1,\\ 1 & u_i\geqslant 1,\end{cases} \quad (22c)$$

where u_i and v_i are the membrane potential and inhibitory currents of the ith node ($i=1,2,...,N$). W_i is the coupling effect from other nodes, $W_i=\sum_{j=1}^{N}M_{ij}u_j(t-\tau_{ij})$, $(i\neq j)$. M_{ij} is the connection intensity and τ_{ij} is the transmission time delay between adjacent nodes. The function $\delta_{1,i}$ means that the external signal is loaded on node 1. Parameters of identical local dynamics for all nodes are $a=0.84; b=0.07; e=0.04$.

A randomly generated network of size $N=30$ is considered (more details are shown in caption of Fig. 6). The schematic diagram of the network is shown in Fig. 6(a), where all the nodes are arranged according to the shortest distance from A. For example, the shortest distances from node $A=1$ to nodes 19 and 26 are $d(1\to 19)=1$ and $d(1\to 26)=4$. The driving intensity on node $A=1$ is set $Q_{DA}=0.1$. According to the reconstruction process, we detect the distance $\tilde{d}_{BA}$ $B=2,3,...,N$ based on the collected data from this excitable network. If the distance of the shortest paths from A to B are accurately reconstructed, the node B is marked with white hollow circle in Fig. 6(a), otherwise it is marked in gray. We can further infer and calculate effective intensity and time delay for paths with correct distance [Eq. (18)]. Figure 6(b) presents the reconstructed intensities $\tilde{I}_{BA}(P)$ against the corresponding structural values $I_{BA}(P)$ for all reconstructed paths together, where the round dots for the shortest paths and the squares for the second shortest paths. The results are mostly quite accurate for the shortest paths, however there are some relatively large errors for the second shortest paths.

To optimize the experimental results, we increase the amount of data and recollect $N_{\rm run}=20$ times different dataset. In each run the system evolves from different initial states and the correlation function in the ith dataset is denoted as $E^{1,d+1}_{AB,i}(t')$. The averaged correlation is thus

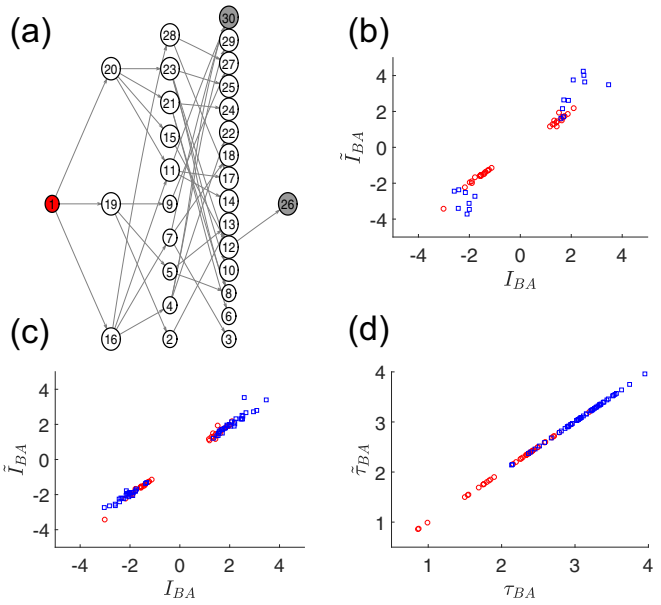

FIG. 6. Reconstruction results of the network of modified FHN (22). The network of size is $N=30$. The connection matrix is randomly generated, with connection probability 0.1. Each connected link is chosen to be positive or negative randomly. Value of time delay and strength of a single link is randomly chosen in range $\tau_{ij}\in(0.5,1)$ and $M_{ij}\in(-1.5,-1)\cup(1,1.5)$. (a) The schematic diagram of the network. All nodes are arranged according to the shortest distance from node A ($A=1$ here and in all the following figures). Nodes marked in gray indicate that their shortest distances are not inferred correctly. (b) The reconstructed interaction intensities $\tilde{I}_{BA}(P)$ based on single dataset plotted versus the structural intensities $I_{BA}(P)$, the red round dots are for the shortest paths and the blue square dots are for the second shortest paths. (c, d) The reconstructed effective intensities and time delays $\tilde{I}_{BA}(P)$ and $\tilde{\tau}_{BA}(P)$ based on 20 different datasets, plotted versus the structural values $I_{BA}(P)$ and $\tau_{BA}(P)$. In panel (c) the reconstructed results are greatly optimized than panel (b).

given by

$$\bar{E}^{1,d+1}_{AB}(t') = \frac{1}{N_{\rm run}}\sum_{i=1}^{\rm Times}E^{1,d+1}_{AB,i}(t'). \quad (23)$$

The new reconstruction results based on the averaged correlation function $\bar{E}^{1,d+1}_{AB}(t')$ are shown in Fig. 6(c), which is greatly improved where all values are almost distributed on the diagonals, i.e., $\tilde{I}_{BA}\approx I_{BA}$. Moreover, we find that more interaction paths, including the paths related to the nodes in gray in Fig. 6(a), are correctly inferred. In addition, the transmission time delays of these paths are perfectly inferred, shown in Fig. 6(d).

We take the path $P: 1\to ...\to 2$ as an example to illustrate the optimization effect of larger dataset. As shown in Fig. 7(a), there are one shortest path with distance $d=2$, and two second shortest paths with distance $d=3$. The shortest path is well reconstructed via $E^{1,3}_{12}(t')$ which is calculated with a single dataset, since a singularity is detected clearly as shown in Fig. 7(b). The multiple second shortest paths cannot be detected, since the characteristic singularities related to the second shortest paths are embedded in

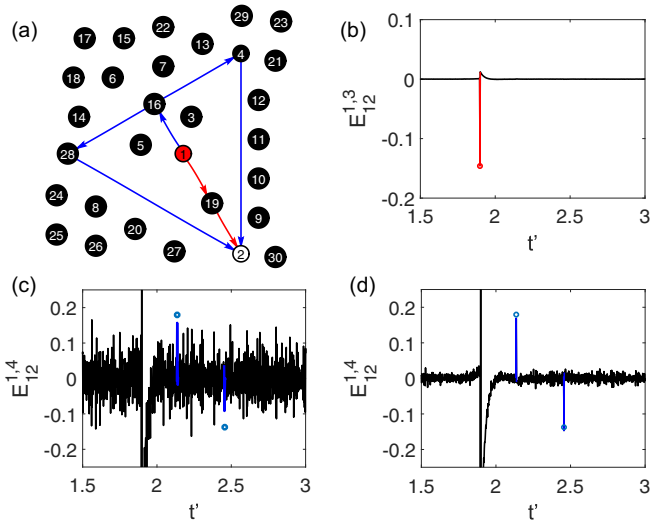

FIG. 7. The illustration of the optimization effect of averaging 20 datasets. The time length of each datasets is $T = 10\,000$. (a) The schematic diagram of paths from $A = 1$ to $B = 2$, where the light red line is the shortest paths and the dark blue lines are the second shortest paths. (b) $E_{12}^{1,3}(t')$ in the range of $t' \in [1.5, 3]$ based on single dataset. A singularity can clearly depict the effective intensity and time delay. (c) $E_{12}^{1,4}(t')$ in the range of $t' \in [1.5, 3]$ based on the same single dataset. The singularities marked are embedded in the perturbations. (d) The averaged correlation function $\tilde{E}_{12}^{1,4}(t')$ via 20 different datasets [Eq. (23)]. Two expected singularities emerge, and the corresponding information of two second shortest paths can be calculated accurately. In panels (b), (c), and (d) and also in all the following figures, circles have the same meanings as in Fig. 3.

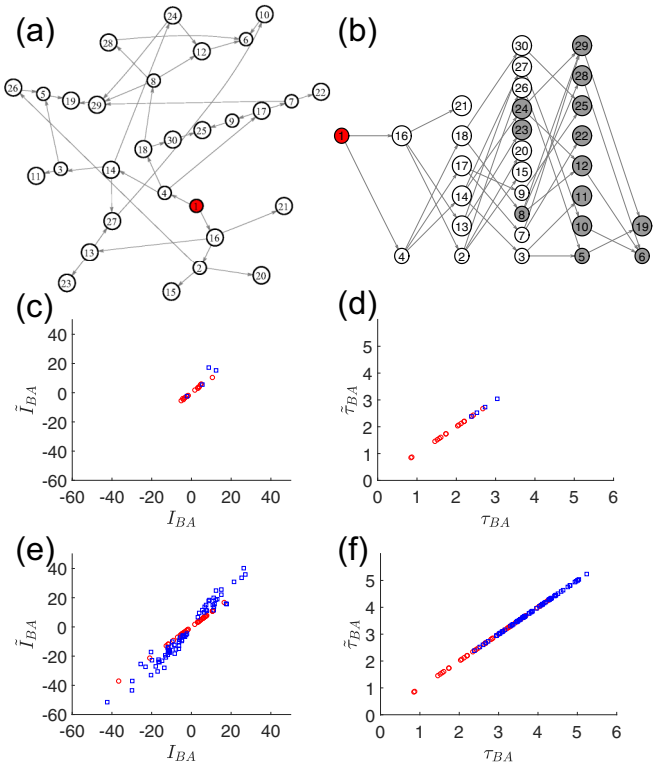

FIG. 8. Reconstruction results of the Lorenz networks Eq. (24). The parameters of local dynamics are $\sigma = 10$, $p = 28$, $q = s = 1$, $a = 1$, $b = 8/3$. (a) The schematic diagram of the network. Each single link is randomly linked in the same way as in Fig. 6. The value of $M_{i,j}$ and $\tau_{i,j}$ are randomly set in range $\tau_{ij} \in (0.5, 1)$ and $M_{ij} \in (-2, -1) \cup (1, 2)$. (b) Rearrangement network. All nodes in the network are rearranged according to the shortest distance from $A = 1$. The nodes in gray indicates that their shortest path cannot be reconstructed even based on 20 different datasets. (c, d) The reconstructed $\tilde{I}_{BA}(P)$ and $\tilde{\tau}_{BA}(P)$ vs the actual structural $I_{BA}(P)$ and $\tau_{BA}(P)$, where the red round dots for the shortest paths and blue square dots for the second shortest paths. The number of reconstructed paths are very limited since the distances of the shortest paths of many nodes cannot be correctly inferred [shown in panel (b)]. (e, f) The improved results of $\tilde{I}_{BA}(P)$ and $\tilde{\tau}_{BA}(P)$ vs the $I_{BA}(P)$ and $\tau_{BA}(P)$. The correlation functions are all improved by the multiple iterative difference approach as Eq. (25). All the paths from A to B are reconstructed by $\Delta^2 \bar{E}_{AB}^{1,d+1}(t')$ successfully based on same data as panels (c, d).

strong noises as shown in Fig. 7(c). However, two singularities emerge in the $\bar{E}_{12}^{1,4}(t')$ averaged by $N_{\text{run}} = 20$ different datasets as shown in Fig. 7(d). Comparing the value of two peaks with the theoretical values denoted by two circles, the two second shortest paths are well reconstructed by a larger dataset.

B. The chaotic Lorenz networks

A coupled Lorenz network with chaotic dynamics is considered, which is given by

$$\begin{aligned}
\dot{x}_i(t) &= \sigma[y_i(t) - x_i(t)] + \sum_{j=1}^{N} M_{ij} x_j(t - \tau_{ij}) + \delta_{i,1} D_1(t), \\
\dot{y}_i(t) &= p x_i(t) - q y_i(t) - s x_i(t) z_i(t), \\
\dot{z}_i(t) &= a x_i(t) y_i(t) - b z_i(t),
\end{aligned} \tag{24}$$

where the driving signal is loaded on the variable x of node $A = 1$ (with intensity $Q_{DA} = 1$). Parameters of identical local dynamics for all nodes are $\sigma = 10$, $p = 28$, $q = s = 1$, $a = 1$, $b = 8/3$. This network is randomly linked in the same way as the excitable networks. The same as Fig. 6, we calculate correlation function $E(t')$ based on single dataset of $T = 10\,000$ for reconstruction of the shortest paths, however for reconstruction of the second shortest paths we calculate the averaged correlation function $\bar{E}(t')$ based on $N_{\text{run}} = 20$ different datasets. Figure 8(a) shows a sketch diagram of the network structure, and all the nodes are arranged by their distances from node $A = 1$ as shown in Fig. 8(b). Distances of almost all nodes with $d_s \geqslant 3$ are failed to be reconstructed and the corresponding nodes B are marked in gray. Of course, inferences of other quantities are failed together. Furthermore, for the nodes near the source node A there are only a few reconstructed paths, which can be shown by the number of points in Fig. 8(c).

These limitations of reconstruction are mainly due to the strong chaotic behavior of single node. The correlations between the two chaotic trajectories changing violently in a wide range are bound to be large, and can be further amplified by the higher-order derivatives of $x_B(t)$. Thus, the singularities related to the shortest and second shortest paths are embedded

in strong disturbances of the correlation. It is inefficient to eliminate them by averaging multiple runs.

Since the dynamic-caused correlation [Eq. (17b)] is irrelevant to a small time gap through t', we develop a multiple iterative difference approach to improve the reconstruction results in our study, which is given by

$$\Delta^1 E(t') = E(t' + \Delta t) - E(t'),$$
$$\Delta^{\nu+1} E(t') = \Delta^{\nu} E(t' + \Delta t) - \Delta^{\nu} E(t'). \tag{25}$$

Either $E_{AB}^{m,n}$ or $\bar{E}_{AB}^{m,n}$ obtained after averaging calculation can combine with this difference operation.

$E(t')$ can be either $E_{AB}^{m,n}$ or the averaged $\bar{E}_{AB}^{m,n}$. Through our test, by using the multiple iterative difference method, all the distances of node B can be determined correctly including those marked in gray in Fig. 8(b). The reconstructed effective intensities and time delays vs the real are plotted in Figs. 8(e) and 8(f), and all the values are around the diagonals which means the detailed information of these paths are reconstructed accurately.

To illustrate the validity of the multiple iterative difference method, we choose the reconstruction of paths from $A = 1$ to $B = 3$ as an example shown in Fig. 9(a), where the distance of shortest path and second shortest path are $d = 3$ and $d = 4$, respectively. A singularity related to the shortest path emerges in the correlation $E_{13}^{1,4}$ in Fig. 9(b), where large waves are around the singularity. The singularities related to the second shortest paths are embedded in the large fluctuations of $\bar{E}_{13}^{1,5}$ and cannot be detected as shown in Fig. 9(d). It is evident that there is a large correlation between the data from any pair of nodes in this chaotic model, which is independent of the driving. In Figs. 9(b) and 9(d) waves caused by the chaotic dynamics fluctuate with large amplitude and change continuously over time, so that the differential operation in Eq. (25) can reduce the waves effectively. For a clear illustration, $\Delta^2 E_{13}^{1,4}$ and $\Delta^2 \bar{E}_{13}^{1,5}$ are calculated as shown in Figs. 9(c) and 9(e), where the singularities related to the shortest and second shortest paths emerge apparently, and the fluctuations in the correlations are well eliminated. The detailed information of the multiple paths can be reconstructed accurately.

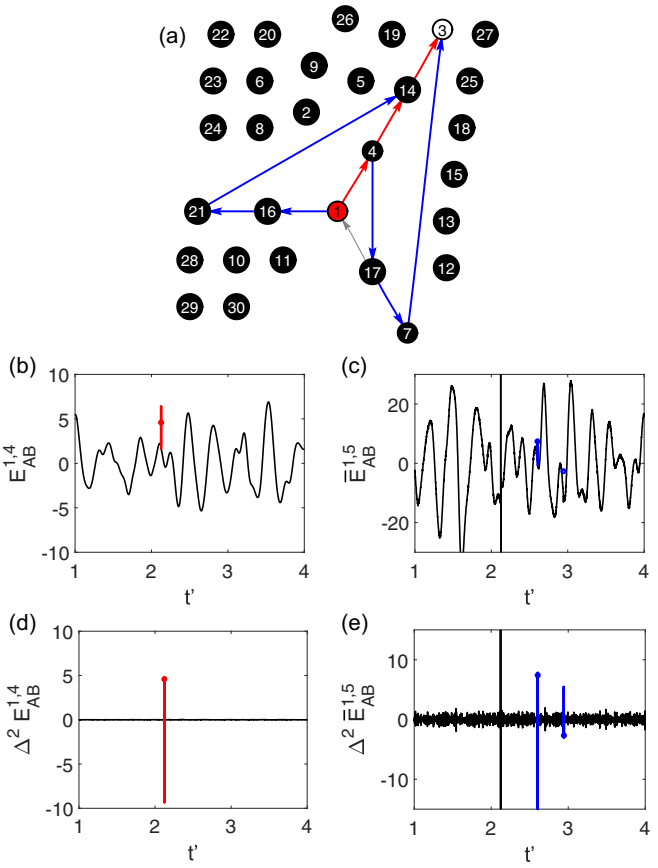

FIG. 9. The illustration of the multiple iterative difference approach. (a) The shortest paths and the second shortest paths from $A = 1$ to $B = 3$ in network of Fig. 8(a). (b) The correlation $E_{AB}^{1,4}(t')$ in the range of $t' \in [1, 4]$. The singularity related to the shortest path is detected, but there are large waves around the singularity, which can cause negative effect on reconstruction. (c) The averaged correlation $\bar{E}_{AB}^{1,5}(t')$ in the range of $t' \in [1, 4]$, where the singularities are embedded in the strong fluctuations. Panels (d) and (e) are calculated using the multiple iterative difference approach Eq. (25). (d) The correlation $\Delta^2 E_{AB}^{1,4}(t')$. The singularity related to the shortest path emerges and the background waves are suppressed considerably. (e) The correlation $\Delta^2 \bar{E}_{AB}^{1,5}(t')$. Like panel (d), two singularities emerge and reconstruct the second shortest paths perfectly.

C. The gene regulatory networks

In the above models, we only consider fixed linear interactions between adjacent nodes. However, many practical systems have variable-dependent and strong nonlinear interactions. For instance, the following model has been widely used to describe the dynamics of gene regulatory network (GRN) [39],

$$\dot{x}_1(t) = \gamma_1 + \sum_{j=1}^{N} \{\Psi_{1j}[x_1(t), x_j(t - \tau_{1j})]\} - x_1 + \delta_{1,i} D_1(t), \tag{26}$$

where

$$\Psi_{ij}(x_i, x_j) = \begin{cases} 0, & \text{no regulation from } j \text{ to } i, \\ (1 - x_i)\dfrac{x_j^{h_{ij}}}{x_j^{h_{ij}} + K_{ij}^{h_{ij}}}, & \text{active regulation}, \\ -x_i \dfrac{x_j^{h_{ij}}}{x_j^{h_{ij}} + K_{ij}^{h_{ij}}}, & \text{repressive regulation}, \end{cases}$$

where x_i represents the expression level of gene i, $x_i \in [0, 1]$, and Ψ_{ij} describes different kinds of regulations between proteins, which are strongly nonlinear. The theoretical structural intensities of any given path should be determined by Eq. (15b) from long time averages of the instantaneous value of interaction function, instead of direct product of link matrices. The intensity can be treated as qualitative judgment of the active or repressive regulation. To maintain the stability of the system with strong nonlinear interactions, a driving signal of weak intensity is applied, i.e., $Q_{DA} = 0.001$.

Reconstruction results of GRN are show in Fig. 10. Almost all the shortest and second shortest paths can be reconstructed accurately, and only the distance from of node 35 marked in gray in Fig. 10(b) is not detected.

The reconstructed effective intensities and time delays versus the theoretical effective intensities and time delays are plotted in Figs. 10(c) and 10(d), most of the values

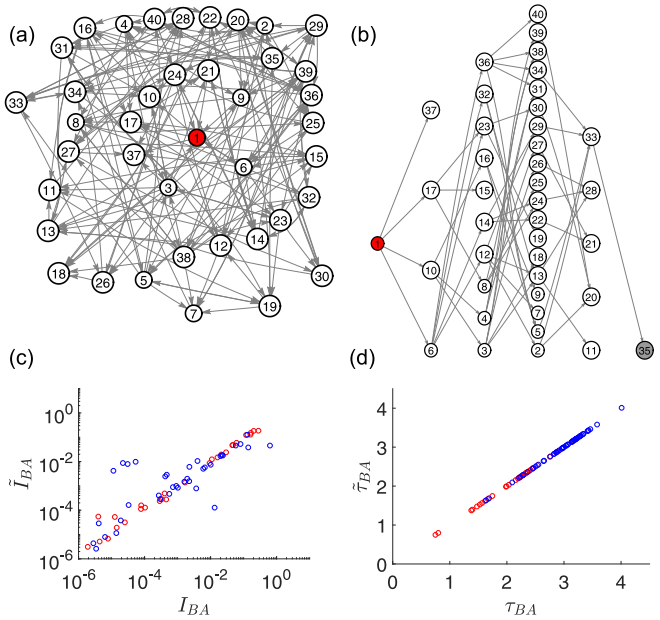

FIG. 10. Reconstruction results of GRN Eq. (26). (a) The schematic diagram of the network structure. We randomly choose $\gamma_i \in (0, 0.1)$, $h_{ij} \in (1, 10)$ and $K_{ij} \in (0, 1)$ within their ranges with equal probabilities. The topological structure with $N = 40$ is randomly constructed. Each node receives inputs from five other nodes. Each input link has half chance to be active or repressive. (b–d) The reconstructed results like in Figs. 6 and 8. Both averaged correlation of $N_{\text{run}} = 10$ and one-order multiple difference method Eq. (25) with correlation are implemented $\Delta^1 \bar{E}_{AB}^{1,d+1}(t')$ for the reconstruction of GRN. In panel (b), almost all the shortest paths are reconstructed correctly, and only the shortest path from node 1 to 35 is not detected (marked in gray). In panels (c) and (d) again $\tilde{I}_{BA}(P)$ and $\tilde{\tau}_{BA}(P)$ are plotted versus the structural values $I_{BA}(P)$ and $\tau_{BA}(P)$. Since the magnitude order of I_{BA} varies greatly, in panel (c) logarithmic diagram are applied for presentation, where the absolute values $|I_{BA}|$ and $|\tilde{I}_{BA}|$ are shown.

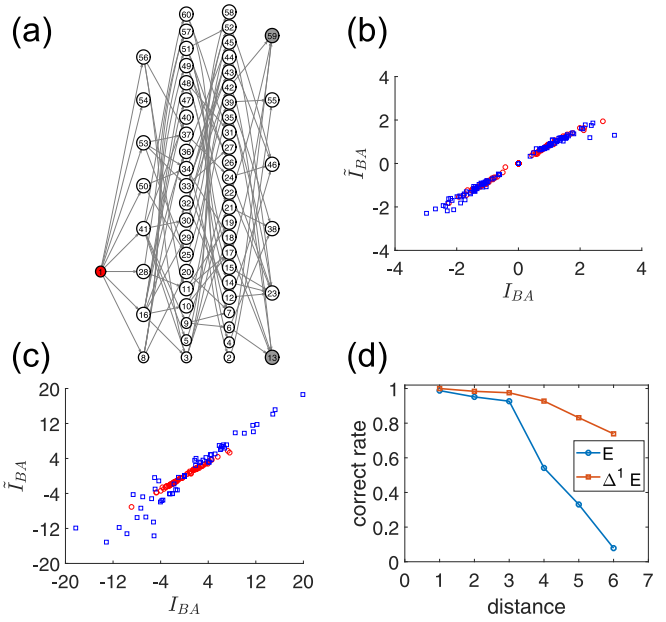

FIG. 11. The reconstructed results of complex networks with larger size. All the networks are constructed in the same way as in Fig. 6 with considerably enlarged size. (a) The schematic diagram of an excitable network with $N = 60$. All the nodes are arranged according to the distance of the shortest path from A. Only a few nodes (colored in gray) are not detected, whose distances of the shortest paths cannot be inferred correctly. (b) The reconstructed effective intensities $\tilde{I}_{AB}(P)$ plotted against $I_{AB}(P)$, of the nodes marked in white hollow circles in panel (a). (c) The same as panel (b) $\tilde{I}_{AB}(P)$ plotted against $I_{AB}(P)$ for a enlarged network with a size $N = 100$. (d) The correct rate of reconstruction results of 6 groups of nodes with different distances $d = 1, 2, 3, 4, 5, 6$. The nodes are randomly chosen from several large excitable networks.

appear around the diagonals, indicating that data-based analyses work well for inferring the nonlinear interactions. Inevitably, there are still certain errors for some reconstructed intensities. Multiplication regulations can cause very small structural interaction intensities and increase difficulties in the reconstruction. In addition, the stochastic driving may cause some changes to the GRN dynamics.

D. The larger excitable networks

We further demonstrate the validity of the deep network reconstruction method in complex systems with much larger size, and take the excitable system Eq. (22) as an example. All parameters are set the same as in Eq. (22) and Fig. 6. In this section, if not specified, time step of the measurement is set $\Delta t = 0.005$. The reconstruction is carried on also in the same way, where the correlation function are averaged by $N_{\text{run}} = 20$ times different dataset, each with $T = 10\,000$. First, we consider a network with network size $N = 60$ with a connection probability $p = 0.09$. The external driving signal is loaded on node $A = 1$, and the rest nodes B ($B \neq A$) are arranged according to the distance of the shortest path from node A in Fig. 11(a). Only a few nodes with large distances are marked in gray whose distances are not distances inferred. As shown in Fig. 11(b) intensities of the shortest and second shortest paths are reconstructed accurately. For path whose distance and effective intensity are correctly inferred, its time delay can be certainly inferred. Second, we create an excitable network with larger size $N = 100$ with a smaller connection probability $p = 0.03$. From statistical analysis, we find that only distances of 5 nodes are failed to be inferred, which are the largest distances in the network. The corresponding $\tilde{I}_{BA}$ versus I_{BA} of the shortest and second shortest paths are shown in Fig. 11(c). In this more sparse and larger network, there are certain more fluctuation errors appear for the second shortest paths due to the strong interactions between multiple paths, since the number of paths with large distances increases a lot.

People usually deal with much larger networks in reality. Finally, we adopt the following strategy to demonstrate the validity of deep reconstruction approach. We build up 10 different excitable networks with random and sparse connections. From node i to another node j, there is a probability of $p = 2/N$ for existence of an one-way link. The network size are chosen randomly $N \in [200, 1000]$. The reconstruction is carried on in the same way, where the correlation function are averaged by $N_{\text{run}} = 2$ of times different dataset, each with $T = 10\,000$. We randomly select not more than

100 different nodes in each network and record reconstruction results of them. According to the distance of the nodes' corresponding shortest path, we group these nodes into 6 groups $d_s = 1, 2, 3, 4, 5, 6$. If the shortest paths of for a given node B are detected by characteristic singularities, we define that this node B is inferred correctly. The correct rates of different groups are shown in Fig. 11(d). The curve underneath is computed by the correlation $\bar{E}$ without the difference operation and the upper is improved by one order difference operation $\Delta^1\bar{E}$ in Eq. (25). As shown in Fig. 11(d), the correct rate decreases as the distance become larger. The correct rate is kept almost equal to 1 for $d \leqslant 3$. However, the correct rate decreases rapidly for further increasing d. With the first-order difference operation, we can infer distances more accurately and rather high successful rates are kept for large d.

In conclusion, it is more difficult to reconstructed the interaction paths with large distance. The deep reconstruction is largely independent of the network size if the shortest distances between any two perceptible nodes is not too large. However, for larger networks, one has to treat cases of longer distances and also more interaction paths between any pair of nodes. Then more data should be measured and the optimization method should be introduced.

V. CONCLUSION

In this paper a new deep reconstruction method is proposed, which is model-free, applicable across system dynamics and network topologies. With an external stochastic driving loaded on a source node A, one can infer representative information about the direct and indirect interactions from A to another node B in the network even when all other nodes are hidden. Rich information can be inferred, including distance from A to B, the effective intensity, and the transmission time delay of multiple paths with different distances. This method can be used to extract rich information because it analyzes not only the correlation of the available data but also the correlation of the multi-order derivatives of the data, which implies the utilization of multidimensional variables and the expansion of the dimensions of the information included in the data. Theoretically, all paths with different distances from A to an arbitrary node B can the reconstructed. In this work, we only consider the inference of the shortest paths and the second shortest paths. The accurate inference of paths with larger distances may need much larger amount of data and other novel optimization operations.

In this paper, we focus on the positive effect of noiselike driving signals, with the premise that one should be able to inject the signal to the source node A. In the future study, exploring the optimal form of the driving signal and developing the corresponding detection technology will be an interesting task in this area. Improving the method based on practicable driving signals may provide a way to obtain broader insights for more powerful functions for network reconstruction. Since in many realistic networks parts of nodes are often not reachable, the approach developed in this paper, reconstructing multiple paths of direct and indirect interactions between two nodes based on information of these two nodes only, is expected to be useful to explore rich information of hidden world of practical networks.

In epilepsy research, detecting system dynamics can tell us that the brain has entered a critical state earlier, so as to better predict seizures and find vulnerable brain regions [5]. The most probable applications of our work may be expected in real neural systems, where in different levels (e.g., single neuron level, local circuits level and functional area level) extremely rich data are available for analyses of network interactions. On the one hand, network structures determine the specific functions of some local circuits, and on the other hand, many network units are unknown or not measurable (hidden) and the approach developed in this paper is hopefully applicable [23,40–44].

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China [Grants No. 11835003 (G.H.), No. T2122016 and No. 11734004 (Y.Y.Mi), No. 11605098 and No. 11975131 (Z.Z.), No. 1190051195 (Y.C.)], National Science and Technology Innovation 2030 Major Program [Grants No. 2021ZD0203700 and No. 2021ZD0203705 (Y.Y. Mi)], and Huawei Technology Co., Ltd (Grant No. YBN2020095106).

X.W. and Y.M. contributed equally to this work.

APPENDIX A: IMPACT OF IRRELEVANT FACTORS

As mentioned above, reconstruction are based on measurable data of a targrt node (B) and external driving (on source node A). Recommending the previous contents of this paper, data sequences of node B and driving on A are shown as follows:

$$x_B(t_1), x_B(t_2), ..., x_B(t_l), ..., x_B(t_L), \quad l = 1, 2, ..., L-1,$$
$$0 < t_{l+1} - t_l = \Delta t \ll 1, T = L\Delta t, \tag{A1}$$

$$D(t_1), D(t_2), ..., D(t_l), ..., D(t_L). \tag{A2}$$

It is known that at least $2 \times L$ numbers are recorded. Sufficient theoretical deduction and numerical examples are given to illustrate the proposed method, where the characteristic discontinuity can be found in the correlation function with t' scanning and covering enough range. However, there are several factors that can always influence the test results, sometimes even greatly. In the Appendix A, effects of these different impacts are discussed.

1. Random deviation caused by noises

First we use noise-induced deviation for nodes with distance $d = 1$ [Eq. (9a)] to give an example:

$$E^1_{\Gamma_B} = \left\langle D_A(t)\frac{\Gamma_B(t+\Delta t) - \Gamma_B(t)}{\Delta t}\right\rangle. \tag{A3}$$

As mentioned above, $\Gamma_B(t)$ denotes the equivalent noise effect accumulated during a very short time. Note that when measuring data with a fixed step Δt at time t_k, $\Gamma_B(t_k)$ represents the equivalent impact of noise on the measurable quantities in a very short time Δt. The value of $\Gamma_B(t_k)$ is an uncertain random

number, but its statistical property can be deduced. According to Eq. (1b), one can get the expectation of $\Gamma_B(t_k)$

$$E(\Gamma_B(t_k)) = \lim_{\Delta t\to 0}\int_{t_k}^{t_k+\Delta t}\Gamma_B(t)dt = 0, \tag{A4a}$$

and the variance

$$V[\Gamma_B(t_k)] = \lim_{\Delta t\to 0}\int_{t_k}^{t_k+\Delta t}[\Gamma_B(t)]^2dt = Q_B\frac{1}{\Delta t}. \tag{A4b}$$

Therefore, $[\Gamma_B(t_k)]$, $k = 1, 2, ..., L$ are a series of independent variables of Gauss distribution. The variance of the noise-induced deviation $E^1_{\Gamma_B}$ can be divided into a driving variance $V(\langle D(t)\rangle)$ and a multiplicative noise variance term

$$\begin{aligned} V\left(E^1_{\Gamma_B}\right) &= V\left(\left\langle D_A(t)\frac{\Gamma_B(t+\Delta t)-\Gamma_B(t)}{\Delta t}\right\rangle\right) \\ &= V(\langle D_A(t)\rangle)V\left(\left\langle\frac{\Gamma_B(t+\Delta t)-\Gamma_B(t)}{\Delta t}\right\rangle\right), \end{aligned}$$

where

$$\begin{aligned} &V\left(\left\langle\frac{\Gamma_B(t+\Delta t)-\Gamma_B(t)}{\Delta t}\right\rangle\right) \\ &= V\left(\frac{1}{L-1}\sum_{k=1}^{L-1}\frac{\Gamma_B(t_k+\Delta t)-\Gamma_B(t_k)}{\Delta t}\right) \\ &= \frac{2Q_B}{(L-1)\Delta t^3} = \frac{2Q_B}{T\Delta t^2}. \end{aligned} \tag{A5}$$

In the above equation, K and H are two independent variables, they have the following properties [for convenience here we use K to represent $\Gamma_B(t_k)$ and H to represent $\Gamma_B(t_h)$]

$$\begin{aligned} D(K+H) &= \langle(K+H-E(K+H))^2\rangle \\ &= E(K^2)+E(H^2)+E^2(K)+E^2(H) \\ &\quad +2E(K)E(H)+2E(K)E(H)-2E^2(K+H) \\ &= E(K^2)-E^2(K)+E(H^2)-E^2(H) \\ &= D(K)+D(H), \end{aligned}$$

and

$$D\left(\frac{1}{L-1}K\right) = \left\langle\left(\frac{1}{L-1}K\right)^2\right\rangle = \left(\frac{1}{L-1}\right)^2 D(K).$$

Similarly, for reconstruction of farther distances d of Eqs. (17) and (18)

$$V\left(\left\langle\Gamma_B^{(d)}\right\rangle\right) = \frac{2^dQ_B}{T\Delta t^{(2d)}}. \tag{A6}$$

The above deduction proves that noises can cause random deviation (as shown in Fig. 5)

$$\begin{aligned} V\left(\Delta E^d_{\Gamma_B}\right) &= V(\langle D_A\rangle)V\left(\left\langle\Gamma_B^{(d)}(t+\Delta t)-\Gamma_B^{(d)}(t)\right\rangle\right) \\ &= V(\langle D_A\rangle)2\frac{2^dQ_B}{T\Delta t^{(2d)}}. \end{aligned} \tag{A7}$$

It is more appropriate and cost-effective to properly increase Δt to weaken noise effects than to increase the time length of data sequences T only. However, increasing Δt excessively may cause confusion of different characteristic discontinuities.

2. Bias caused by system dynamics

First, for the reconstruction of distance $d = 1$. In Eq. (9b) the value of $\Phi_i[\vec{X}_B(t)]$ is independent of the time delays. For convenience, in this part parameters of time delay are omitted ($\tau_{ij} = 0$, for arbitrary i, j). So the variables can be simplified as $\vec{X}_B(t) \to x_B(t)$. For convenience, Eq. (9b) is also simplified

$$\begin{aligned} &\sum_{i=1}^{N}\frac{\partial\Phi_B[x_B(t)]}{\partial x_i(t)}\Phi_i(x_B(t)) \to F_B^1[x_B(t)], \\ &E^1_{\Phi_B} = \left\langle D_A(t)F_B^1[x_B(t)]\right\rangle, \end{aligned} \tag{A8}$$

where $F_B^1(x_B)$ is a continuous and differentiable function whose variable $x_B(t)$ changes over time t. According to the property that the driving is independent of the system dynamics (D_A is independent of F_B^1), the first-order differential of $E^1_{\Phi_B}$ is given as

$$\begin{aligned} \Delta E^1_{\Phi_B} &= \left\langle D_A(t)F_B^1[x_B(t+\Delta t)]\right\rangle - \left\langle D_A(t)F_B^1(x_B(t))\right\rangle \\ &= \left\langle D_A(t)\left(F_B^1[x_B(t+\Delta t)]-F_B^1(x_B(t))\right)\right\rangle \\ &= \langle D_A(t)\rangle\left\langle F_B^1[x_B(t+\Delta t)]-F_B^1(x_B(t))\right\rangle. \end{aligned} \tag{A9}$$

Based on the Lagrange Mean Value Theorem,

$$\begin{aligned} &\left\langle F_B^1[x_B(t+\Delta t)]-F_B^1[x_B(t)]\right\rangle \\ &= \left\langle\dot{F}_B^1[x_B(t+\theta\Delta t)][x_B(t+\Delta t)-x_B(t)]\right\rangle \\ &\approx \left\langle\dot{F}_B^1[x_B(t+\theta\Delta t)]\right\rangle\langle\dot{x}_B(t)\rangle\Delta t \\ &= \left\langle\dot{F}_B^1[x_B(t)]\right\rangle\langle\Phi_B[x_B(t)]\rangle\Delta t, \end{aligned} \tag{A10}$$

where $0 < \theta < 1$ and $\theta\Delta t$ is omitted again in average calculation. In conclusion, for reconstruction of distances $d = 1$,

$$\Delta E^1_{\Phi_B} = \langle D_A(t)\rangle\left\langle\dot{F}_B^1(x_B(t))\right\rangle\langle\Phi_B(x_B(t)\rangle\Delta t. \tag{A11}$$

Then, for reconstruction of distances 2

$$\begin{aligned} &\frac{d}{dt}\left(\sum_{i=1}^{N}\frac{\partial\Phi_B[x_B(t)]}{\partial x_i(t)}\Phi_i[x_B(t)]\right) \to F_B^2[x_B(t)], \\ &E^2_{\Phi_B} = \left\langle D_A(t)F_B^2[x_B(t)]\right\rangle, \end{aligned} \tag{A12}$$

and

$$\Delta E^2_{\Phi_B} = \langle D_A(t)\rangle\left\langle\dot{F}_B^2[x_B(t)]\right\rangle\langle\Phi_B[x_B(t)]\rangle\Delta t. \tag{A13}$$

Similarly, for reconstruction of farther distances d

$$\begin{aligned} &\frac{d^{d-1}}{dt^{d-1}}\left(\sum_{i=1}^{N}\frac{\partial\Phi_B[x_B(t)]}{\partial x_i(t)}\Phi_i[x_B(t)]\right) \to F_B^2[x_B(t)], \\ &E^d_{\Phi_B} = \left\langle D_A(t)F_B^d[x_B(t)]\right\rangle, \end{aligned} \tag{A14}$$

and

$$\Delta E^d_{\Phi_B} = \langle D_A(t)\rangle\left\langle\dot{F}_B^d[x_B(t)]\right\rangle\langle\Phi_B[x_B(t)]\rangle\Delta t. \tag{A15}$$

Under ideal conditions, i.e., $T \to +\infty$ and so $\langle D_A(t)\rangle \to 0$, system dynamics contribute no variable component

$\Delta E^d_{\Phi_B} \to 0$ in the correlation function. Under realistic conditions with finite data volume, $\langle D_A(t)\rangle = \sum_{i=1}^{L} D_A(t_L) \neq 0$. So the correlation function $E^{1,d+1}_{AB}(t')$ changes with t' and looks like smooth fluctuations. However, we can effectively weaken the fluctuations by multiple difference approach Eq. (25) because $\Delta E_{\Phi_B} \to O(\Delta t)$.

Various system dynamics may be so complicated that we can only give a qualitative estimation roughly about the size of the bias caused by dynamics. The size of the bias depends on the system. In the same system the larger the path distance is, the larger the bias is. If the bias exceeds the amplitude of the characteristic discontinuity, it may cause some difficulties. The high-order derivation of data can be a simple metric for $E^d_{\Phi_B}$:

$$\left\langle E^d_{\Phi_B} \right\rangle \propto \left\langle x_i^{(d)}(t)\right\rangle or \left\langle \left| x_i^{(d)}(t)\right|\right\rangle. \tag{A16}$$

In summary, system dynamics can cause bias in $E_{AB}(t')$, which changes continuously as t'. The bias changes randomly with an averaging rate of an equivalent magnitude which can grow larger when reconstructing farther distances. So we choose a nonperiodic driving signal with small Δt in this work, and a differential approach is proposed to effectively eliminate dynamic-induced bias.

APPENDIX B: IMAGE DISCREPANCY CAUSED BY NUMERICAL SIMULATION INTERVAL

As is mentioned in the caption of Fig. 5(d), unexpectedly, there are more than one discontinuous points in correlation function. This is mainly caused by the rearrangement operation and systematic feature of numerical simulation.

In this paper, we use the Stochastic Runge-Kutta algorithms [45] for numerical simulation. The numerical simulation is performed with interval dt. In the numerical simulation process of variable data, we assume every numbers obtained by the numerical simulation is measured and recorded, i.e., $dt = \Delta t$.

From the causal relations, driving at time t can affect the source node with a short delay but cannot affect source node at the same time:

$$\langle x_A(t) D_A(t)\rangle = 0, \tag{B1a}$$

$$\langle x_A(t+\Delta t) D_A(t)\rangle \neq 0. \tag{B1b}$$

Besides, for general systems with dissipative property short-term correlation decays fast ($k \geqslant 2$)

$$\langle x_A(t+k\Delta t) D_A(t)\rangle = 0. \tag{B2}$$

So there is a discontinuity comes from the correlation about $x_A(t+\Delta t) - x_A(t)$ and $D_A(t)$, which is also called that driving component in $x_A(t+\Delta t) - x_A(t)$ comes from $D_A(t)$.

$$D_A(t) \to x_A(t+\Delta t) - x_A(t). \tag{B3}$$

Driving on node A can also affect other nodes along interaction paths. For example, if there is a path $A \to B \to C \to$ ..., then the distance of B from A is $d = 1$, the distance of C from A is $d = 2$, and so on. This constraint can be extended to node B with a time delay t_{BA} from A. (For convenience, in the following we talk about situation with zero time delay $t_{BA} = 0$, $t_{CB} = 0$,...) From the above, driving component in $x_B(t+\Delta t) - x_B(t)$ comes from $x_A(t)$:

$$\begin{aligned} x_A(t) &\to x_B(t+\Delta t) - x_B(t) \\ &\quad or \\ x_A(t) &\to x_B^{(1)}(t). \end{aligned} \tag{B4}$$

Similarly, the driving component is propagated

$$\begin{aligned} x_B(t) &\to x_C(t+\Delta t) - x_C(t) \\ &\quad \text{or} \\ x_B(t) &\to x_C^{(1)}(t). \end{aligned} \tag{B5}$$

When we calculate $E^{1,2}_{AB}(t')$ based on data of node B, one can find that $x_B^{(2)}(t)$ only contains driving component of $D_A(t)$ by backstepping

$$\begin{aligned} x_B^{(2)}(t) &= \frac{1}{\Delta t}\left[x_B^{(1)}(t+\Delta t) - x_B^{(1)}(t)\right] \\ &\leftarrow [x_A(t+\Delta t) - x_A(t)] \leftarrow D_A(t). \end{aligned} \tag{B6}$$

When we calculate $E^{1,3}_{AC}(t')$ based on data of node C, one can also find that $x_C^{(3)}(t)$ only contains driving component of $D_A(t)$:

$$\begin{aligned} x_C^{(3)}(t) &= \frac{1}{\Delta t}\left[x_C^{(2)}(t+\Delta t) - x_C^{(2)}(t)\right] \\ &= \frac{1}{\Delta t^2}\left[x_C^{(1)}(t+2\Delta t) - x_C^{(1)}(t+\Delta t)\right] \\ &\quad - \frac{1}{\Delta t^2}\left\{\left[x_C^{(1)}(t+\Delta t) - x_C^{(1)}(t)\right]\right\} \\ &\quad \leftarrow \left[x_B^{(1)}(t+\Delta t) - x_B^{(1)}(t)\right] \leftarrow D_A(t). \end{aligned} \tag{B7}$$

In a similar way, there is a one-to-one correspondence between $x_A^{(1)}(t)$, $x_B^{(2)}(t)$, $x_C^{(3)}(t)$,... and $D(t)$. Generally the discontinuity in correlation function can only come from one point as shown in most of figures about E_{AB}.

However, it is complex if the original data are rearranged by Eq. (21) as follows:

$$\begin{aligned} &x_i'(t_1) = x_i(t_k),\ x_i'(t_2) = x_i(t_{2k}),\ ...,\ x_i'(t_{L/k}) = x_i(t_L), \\ &D'(t_1) = \sum_{l=1}^{k} D(t_l),\ D'(t_2) = \sum_{l=k+1}^{2k} D(t_l),\ ..... \\ &D'(t_{L/k}) = \sum_{l=L-k+1}^{L} D(t_l). \end{aligned} \tag{B8}$$

For example, if we set $k = 2$ and initial time as $t_0 = t$

$$\begin{aligned} &x_i'(t_0) = x_i(t),\ x_i'(t_1) = x_i(t+2\Delta t),\ x_i'(t_2) = x_i(t+4\Delta t),\ ... \\ &D'(t_1) = D(t) + D(t+\Delta t), \\ &D'(t_2) = D(t+2\Delta t) + D(t+3\Delta t), \\ &D'(t_3) = D(t+4\Delta t) + D(t+5\Delta t),\ ..., \end{aligned} \tag{B9}$$

then we have

$$\begin{aligned} x_B^{(2)}(t) &= \frac{1}{2\Delta t}\left\{\left[x_B^{(1)}(t+2\Delta t) - x_B^{(1)}(t)\right]\right\} \\ &= \frac{1}{(2\Delta t)^2}\{[x_B(t+4\Delta t) - x_B(t+2\Delta t)]\} \end{aligned}$$

$$
\begin{aligned}
&-\frac{1}{(2\Delta t)^2}\{[x_B(t+2\Delta t)-x_B(t)]\}\\
&\leftarrow [x_A(t+3\Delta t)+x_A(t+2\Delta t)-x_A(t+\Delta t)-x_A(t)]\\
&\leftarrow D_A(t+2\Delta t)+D_A(t+\Delta t)+D_A(t+\Delta t)+D_A(t).
\end{aligned}
\tag{B10}
$$

This means that rearranged $x_B^{(2)}(t)$ contains driving component of $D_A(t+2\Delta t)$, $D_A(t+\Delta t)$, $D_A(t)$. There is a one-many correspondence between $x_B^{(2)}(t)$ and $D_A(t+2\Delta t)$, $D_A(t+\Delta t)$, $D_A(t)$. By backstepping we find that

$$
\langle x_B^{(2)}(t)[D(t)+D(t+\Delta t)]\rangle \neq 0. \tag{B11}
$$

Similarly,

$$
x_B^{(2)}(t-2\Delta t) \leftarrow D_A(t)+D_A(t-\Delta t)+D_A(t-2\Delta t), \tag{B12}
$$

and

$$
\langle x_B^{(2)}(t-2\Delta t)[D(t)+D(t+\Delta t)]\rangle \neq 0 \tag{B13}
$$

The driving component in two data points $x_B^{(2)}(t)$ and $x_B^{(2)}(t-2\Delta t)$ can be detected by $D'_A(t_1)=D_A(t)+D_A(t+\Delta t)$. In this way, when we calculate $E_{AB}^{1,2}(t')$ based on rearranged data of node B, in the correlation function the discontinuity consists of two points close to each other.

Similarly, for node C of distance $d=3$,

$$
\begin{aligned}
&\langle x_C^{(3)}(t)[D(t)+D(t+\Delta t)]\rangle \neq 0,\\
&\langle x_C^{(3)}(t-2\Delta t)[D(t)+D(t+\Delta t)]\rangle \neq 0,\\
&\langle x_C^{(3)}(t-4\Delta t)[D(t)+D(t+\Delta t)]\rangle \neq 0.
\end{aligned}
\tag{B14}
$$

By backstepping there are driving components in three data points $x_C^{(3)}(t)$, $x_C^{(3)}(t-2\Delta t)$, and $x_C^{(3)}(t-4\Delta t)$. So in the correlation function $E_{AC}^{1,3}(t')$ the discontinuity consists of three points.

By analyzing high-order correlation with rearranged data, and so on, for node with distance d, in the correlation function the discontinuity will consists of $d+1$ points. It's complicated to obtain the value of each point. However, the total driving component are fixed. In conditions when data are not rearranged, we only calculate discontinuity value based on one point. In conditions when data are rearranged, we can calculate the sum of all mutation values. In Fig. 5(d), there are three section of mutations between four points. From the point (2.16,0.04136) to (2.2,1.1519), there is an uplift mutation. From the point (2.2,1.1519) to (2.28, −1.2094), there is a drop mutation. By subtraction of the two minimums in two sides from the positive maximum 1.1519, the effective intensity is

$$
\begin{aligned}
\tilde{I}_{BA}(P_2) &= 1.1519-0.04136+1.1519-(-1.2094)\\
&= 3.4718.
\end{aligned}
$$

[1] D. J. Watts and S. H. Strogatz, Collective dynamics of "small-world" networks, Nature (London) **393**, 440 (1998).

[2] A.-L. Barabási and R. Albert, Emergence of scaling in random networks, Science **286**, 509 (1999).

[3] R. Albert and A.-L. Barabási, Statistical mechanics of complex networks, Rev. Mod. Phys. **74**, 47 (2002).

[4] A.-L. Barabási and Z. N. Oltvai, Network biology: Understanding the cell's functional organization, Nat. Rev. Genet. **5**, 101 (2004).

[5] A. Li, C. Huynh, Z. Fitzgerald, I. Cajigas, D. Brusko, J. Jagid, A. Claudio, A. Kanner, J. Hopp, S. Chen, J. Haagensen, E. Johnson, W. Anderson, N. Crone, S. Inati, K. Zaghloul, J. Bulacio, J. Gonzalez-Martinez, and S. Sarma, Neural fragility as an EEG marker of the seizure onset zone, Nat. Neurosci. **24**, 1465 (2021).

[6] S. H. Strogatz, Exploring complex networks, Nature (London) **410**, 268 (2001).

[7] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Complex networks: Structure and dynamics, Phys. Rep. **424**, 175 (2006).

[8] E. Ott, J. H. Platig, T. M. Antonsen, and M. Girvan, Echo phenomena in large systems of coupled oscillators, Chaos: Interdiscip. J. Nonlin. Sci. **18**, 037115 (2008).

[9] R. Phogat, S. Sinha, and P. Parmananda, Echo in complex networks, Phys. Rev. E **101**, 022216 (2020).

[10] M. Timme and J. Casadiego, Revealing networks from dynamics: An introduction, J. Phys. A: Math. Theor. **47**, 343001 (2014).

[11] T. S. Gardner, D. di Bernardo, D. Lorenz, and J. J. Collins, Inferring genetic networks and identifying compound mode of action via expression profiling, Science **301**, 102 (2003).

[12] M. Timme, Revealing Network Connectivity from Response Dynamics, Phys. Rev. Lett. **98**, 224101 (2007).

[13] W.-X. Wang, Q. Chen, L. Huang, Y.-C. Lai, and Mary Ann F. Harrison, Scaling of noisy fluctuations in complex networks and applications to network prediction, Phys. Rev. E **80**, 016116 (2009).

[14] J. Ren, W.-X. Wang, B. Li, and Y.-C. Lai, Noise Bridges Dynamical Correlation and Topology in Coupled Oscillator Networks, Phys. Rev. Lett. **104**, 058701 (2010).

[15] E. S. C. Ching, P.-Y. Lai, and C. Y. Leung, Extracting connectivity from dynamics of networks with uniform bidirectional coupling, Phys. Rev. E **88**, 042817 (2013).

[16] E. S. C. Ching, P.-Y. Lai, and C. Y. Leung, Reconstructing weighted networks from dynamics, Phys. Rev. E **91**, 030801(R) (2015).

[17] Z. Zhang, Z. Zheng, H. Niu, Y. Mi, S. Wu, and G. Hu, Solving the inverse problem of noise-driven dynamic networks, Phys. Rev. E **91**, 012814 (2015).

[18] X. Wang, Z. Zhang, H. Li, Y. Chen, Y. Mi, and G. Hu, Exploring node interaction relationship in complex networks by using high-frequency signal injection, Phys. Rev. E **103**, 022317 (2021).

[19] U. Parlitz, Estimating Model Parameters from Time Series by Autosynchronization, Phys. Rev. Lett. **76**, 1232 (1996).

[20] M. K. S. Yeung, J. Tegnér, and J. J. Collins, Reverse engineering gene networks using singular value decomposition and robust regression, Proc. Natl. Acad. Sci. USA **99**, 6163 (2002).
[21] I. T. Tokuda, S. Jain, I. Z. Kiss, and J. L. Hudson, Inferring Phase Equations from Multivariate Time Series, Phys. Rev. Lett. **99**, 064101 (2007).
[22] W.-X. Wang, R. Yang, Y.-C. Lai, V. Kovanis, and C. Grebogi, Predicting Catastrophes in Nonlinear Dynamical Systems by Compressive Sensing, Phys. Rev. Lett. **106**, 154101 (2011).
[23] A. Das and I. R. Fiete, Systematic errors in connectivity inferred from activity in strongly recurrent networks, Nat. Neurosci. **23**, 1286 (2020).
[24] D. Marbach, R. J. Prill, T. Schaffter, C. Mattiussi, D. Floreano, and G. Stolovitzky, Revealing strengths and weaknesses of methods for gene network inference, Proc. Natl. Acad. Sci. USA **107**, 6286 (2010).
[25] Y.-C. Lai, Controlling complex, nonlinear dynamical networks, Natl. Sci. Rev. **1**, 339 (2014).
[26] R. Hegger, M. J. Bünner, H. Kantz, and A. Giaquinta, Identifying and Modeling Delay Feedback Systems, Phys. Rev. Lett. **81**, 558 (1998).
[27] V. Ponomarenko and M. Prokhorov, Recovery of equations of coupled time-delay systems from time series, Tech. Phys. Lett. **31**, 64 (2005).
[28] I. V. Sysoev, V. I. Ponomarenko, D. D. Kulminskiy, and M. D. Prokhorov, Recovery of couplings and parameters of elements in networks of time-delay systems from time series, Phys. Rev. E **94**, 052207 (2016).
[29] R.-Q. Su, W.-X. Wang, and Y.-C. Lai, Detecting hidden nodes in complex networks from time series, Phys. Rev. E **85**, 065201(R) (2012).
[30] R.-Q. Su, Y.-C. Lai, X. Wang, and Y. Do, Uncovering hidden nodes in complex networks in the presence of noise, Sci. Rep. **4**, 3944 (2014).
[31] E. S. C. Ching and P. H. Tam, Effects of hidden nodes on the reconstruction of bidirectional networks, Phys. Rev. E **98**, 062318 (2018).
[32] H. Haehne, J. Casadiego, J. Peinke, and M. Timme, Detecting Hidden Units and Network Size from Perceptible Dynamics, Phys. Rev. Lett. **122**, 158301 (2019).
[33] M. E. J. Newman, Network structure from rich but noisy data, Nat. Phys. **14**, 542 (2018).
[34] R. Benzi, A. Sutera, and A. Vulpiani, The mechanism of stochastic resonance, J. Phys. A: Math. Gen. **14**, L453 (1981).
[35] J. Collins, C. Chow, A. Capela, and T. Imhoff, Aperiodic stochastic resonance, Phys. Rev. E **54**, 5575 (1996).
[36] Hu Gang, T. Ditzinger, C. Z. Ning, and H. Haken, Stochastic Resonance without External Periodic Force, Phys. Rev. Lett. **71**, 807 (1993).
[37] A. S. Pikovsky and J. Kurths, Coherence Resonance in a Noise-Driven Excitable System, Phys. Rev. Lett. **78**, 775 (1997).
[38] M. Bär and M. Eiswirth, Turbulence due to spiral breakup in a continuous excitable medium, Phys. Rev. E **48**, R1635 (1993).
[39] T. Y.-C. Tsai, Y. S. Choi, W. Ma, J. R. Pomerening, C. Tang, and J. E. Ferrell, Robust, tunable biological oscillations from interlinked positive and negative feedback loops, Science **321**, 126 (2008).
[40] E. Schneidman, M. J. Berry, R. Segev, and W. Bialek, Weak pairwise correlations imply strongly correlated network states in a neural population, Nature **440**, 1007 (2006).
[41] J. W. Pillow, J. Shlens, L. Paninski, A. Sher, A. M. Litke, E. J. Chichilnisky, and E. P. Simoncelli, Spatiotemporal correlations and visual signalling in a complete neuronal population, Nature **454**, 995 (2008).
[42] K. J. Friston, Functional and effective connectivity: A review, Brain Connect. **1**, 13 (2011).
[43] A. Pakman, J. Huggins, C. Smith, and L. Paninski, Fast state-space methods for inferring dendritic synaptic connectivity, J. Comput. Neurosci. **36**, 415 (2014).
[44] A. E. Urai, B. Doiron, A. M. Leifer, and A. K. Churchland, Large-scale neural recordings call for new insights to link brain and behavior, Nat. Neurosci. **25**, 11 (2022).
[45] R. L. Honeycutt, Stochastic Runge-Kutta algorithms. I. White noise, Phys. Rev. A **45**, 600 (1992).