PAPER • OPEN ACCESS

Reconstructing dynamics of complex systems from noisy time series with hidden variables

To cite this article: Zishuo Yan *et al* 2023 *New J. Phys.* **25** 083011

View the article online for updates and enhancements.

New Journal of Physics

The open access journal at the forefront of physics

PAPER

CrossMark

Reconstructing dynamics of complex systems from noisy time series with hidden variables

Zishuo Yan[1]**, Lili Gui**[2]**, Kun Xu**[2] **and Yueheng Lan**[1,2,*]

[1] School of Science, Beijing University of Posts and Telecommunications, Beijing 100876, People's Republic of China
[2] State Key Lab of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing 100876, People's Republic of China
[*] Author to whom any correspondence should be addressed.

E-mail: lanyh@bupt.edu.cn

Abstract

Reconstructing the equation of motion and thus the network topology of a system from time series is a very important problem. Although many powerful methods have been developed, it remains a great challenge to deal with systems in high dimensions with partial knowledge of the states. In this paper, we propose a new framework based on a well-designed cost functional, the minimization of which transforms the determination of both the unknown parameters and the unknown state evolution into parameter learning. This method can be conveniently used to reconstruct structures and dynamics of complex networks, even in the presence of noisy disturbances or for intricate parameter dependence. As a demonstration, we successfully apply it to the reconstruction of different dynamics on complex networks such as coupled Lorenz oscillators, neuronal networks, phase oscillators and gene regulation, from only a partial measurement of the node behavior. The simplicity and efficiency of the new framework makes it a powerful alternative to recover system dynamics even in high dimensions, which expects diverse applications in real-world reconstruction.

1. Introduction

Complex interacting systems are ubiquitous and essential for the modern world which include both the social ones: financial systems, power systems, transportation systems, the internet, and the natural ones: neural networks, gene regulation systems and the earth system [1–12]. Diverse and intriguing behaviour is seen arising, out of the complexity of dynamics of individual agents constituting the network, or more often, of the interactions between them. Although more and more data are now measured in these systems, it remains a great challenge to reconstruct network structures and dynamics because a direct measurement of them is rarely possible [13–18], which nevertheless are critical for an effective description, analysis, prediction and control of system behaviour.

System reconstruction is an inverse problem that has been studied extensively in different contexts at various levels [19–23]. With the time-delay embedding technique [24], an equivalent phase space may be constructed and many dynamical features could be computed but interaction rules between different coordinates need further exploration. Another set of convenient tools are based on information theoretics [25–28], which quickly deduce the network topology from time series but the exact strengths of connections are hard to retrieve. Moreover, with these techniques, a computation of direct interaction links could be very cumbersome in complex networks [12]. To obtain equations of motion, model-based methods are used to determine unknown parameters in the pre-assumed dynamical models, such as differential equations or discrete maps. However, many of these methods only apply to low-dimensional systems [29, 30]. To conveniently reconstruct networked systems, methods based on cybernetics were designed [31–37], so that the auxiliary response network synchronizes with the studied network (driving network) via the measured time series, if the network parameters are correctly estimated. However, when applying these

methods, we should know the node dynamical equations and measured the states of all nodes with no interference, which may not be possible in practice.

Almost invariably, noise is an unavoidable factor in almost all practical reconstruction, which constitutes a major difficulty on various occasions. Distinct methods have been developed to remove or utilize noise in a time series to recover the interaction patterns based on noise-induced dynamics or correlations [38–42], in which either a small noise limit is assumed or priori smoothing of the data should be performed. To utilize all the data to suppress fluctuations, Wang *et al* [43] proposed a globally invariant local fitting method, which is capable of processing data with moderate or strong noise. For high-dimensional systems, if the structure of the complex network is known, Gao and Yan [44] developed a two-stage approach (first stage, global regression; second stage, local fine-tuning) for autonomous inference of node and interaction dynamics of complex systems, which, however, needs knowledge of the network topology and the state of all nodes.

Lack of measurement of all variables is surely another major problem for dynamics reconstruction, which requires the recovery of not only unknown system parameters but also the time course of unmeasured variables—the so-called hidden variables. In 1990, Breeden and Hübler [45] discussed how to reconstruct system equations in the presence of hidden variables. Their method is based on the flow algorithm developed by Cremers and Hübler [46], which is similar to that of James *et al* [47]. It can be applied to systems with one or more such hidden variables, provided that the form of the reconstruction function is known. In a different trial, Gouesbet [48, 49] introduced the concept of standard system to numerically evaluate a set of unknown constants related to the target vector field but high-order derivatives are needed which may be hard to compute in the presence of noise. Wang *et al* [50–54] developed a completely data-driven method by using compressed sensing, which effectively infers the existence and location of hidden nodes in a network based on anomalous prediction errors. However, the determination of the coupling strength requires further investigation. With an information theoretic approach, Wu *et al* [55] used the piecewise approximation to carry out a partial Granger causality test (Guo *et al* [56]) to detect the hidden variables with minor impact. The method of high-order cross correlation (HOCC) [57–59] is proposed as a general technique to reconstruct networked systems driven by noise and with hidden variables. However, the requirement of high-order derivatives could easily amplify noise, which fails the reconstruction if the noise is not small. Similar methods were developed and explained by Ching and Tam [39, 40, 60] based on observable covariances.

In all, the presence of hidden variables brings considerable challenges to the network dynamics reconstruction, especially when the system dimension is high and the dynamics is contaminated with noise. Here we propose a new framework to do the reconstruction based on a cost functional, which infers both the unknown parameters and unmeasured time series of hidden variables simultaneously. A gradient descent algorithm is employed and leads to a set of ordinary differential equations (ODEs). The measured time series are used as inputs which drive these ODEs such that the cost functional approaches to zero. Thus it is possible to recover the parameters and trajectories in real time with online input. Even in the presence of noise and in high dimensions, the proposed reconstruction scheme is still valid. As expected, the more the measured data, the less the derived equations. Because of the evolution character of the method, different from most of the current existing methods, the unknown parameters could enter the equations in a nonlinear way. Successful application of the method to typical examples of network models demonstrated its efficiency and great potential in dynamics determination in complex systems.
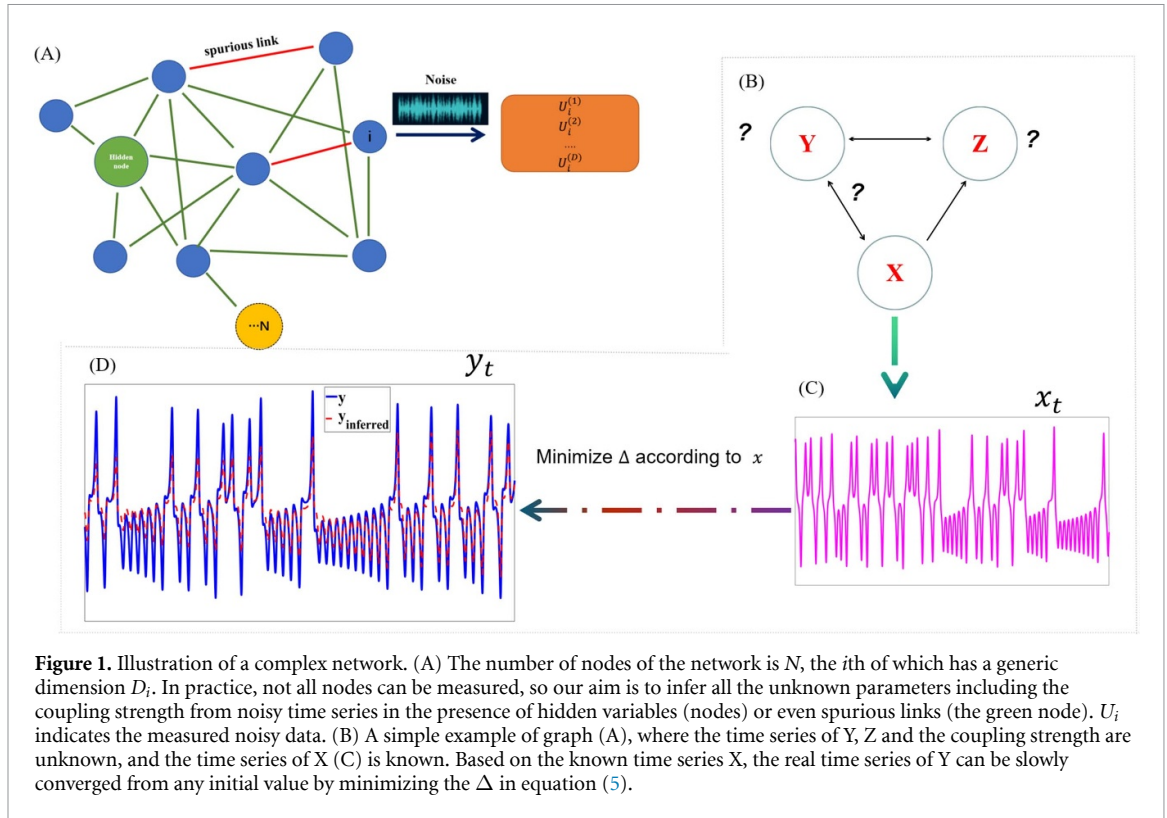
The rest of the paper is organized as follows. The main ideas of our approach are presented in detail in section 2. In section 3, we first apply the method to the well-known Lorenz system driven by white noise, and check the impact of noise intensity and parameter selection. Then, its validity is further demonstrated in application to networks with different local dynamics: the Lorenz oscillators or the FitzHugh–Nagumo(FHN) neural network models. Finally, a discussion of hidden nodes and applicability of the method is made with an example model of coupled Kuramoto oscillators. In the final section, we conclude this paper and point out possible future directions.

2. Results

2.1. Theory
We consider a system consisting of *N* nodes, where the dynamics of each node is determined by a set of differential equations with interactions from other nodes that form a network, given by:

$$\dot{\mathbf{X}}_i = \mathbf{F}_i(\mathbf{X}, \mathbf{p}) = \mathbf{f}_i(\mathbf{X}_i, \mathbf{B}_i) + \sum_{j=1}^{N} A_{ij}\mathbf{g}_{ij}(\mathbf{X}_i, \mathbf{X}_j) \tag{1}$$

Figure 1. Illustration of a complex network. (A) The number of nodes of the network is N, the ith of which has a generic dimension D_i. In practice, not all nodes can be measured, so our aim is to infer all the unknown parameters including the coupling strength from noisy time series in the presence of hidden variables (nodes) or even spurious links (the green node). U_i indicates the measured noisy data. (B) A simple example of graph (A), where the time series of Y, Z and the coupling strength are unknown, and the time series of X (C) is known. Based on the known time series X, the real time series of Y can be slowly converged from any initial value by minimizing the Δ in equation (5).

where $i,j \in 1,2,3\ldots,N$ and $\mathbf{X}_i = \left[X_i^{(1)}, X_i^{(2)}, \ldots, X_i^{(D)}\right] \in \mathbb{R}^D$ is the state vector of node i. The vector $\mathbf{F}_i$ dictates the motion of the ith node with $\mathbf{f}_i, \mathbf{g}_{ij}$ being smooth functions. The parameters $\mathbf{p}$ include the local ones $\mathbf{B}_i$ and the coupling strength A_{ij}. The functional form $\mathbf{f}_i$ and $\mathbf{g}_{ij}$ are assumed known but the parameters $\mathbf{p}$ have to be determined from the given time series.

In the following, we present a method for reconstructing the network connection and equation parameters with only a partial observation of the system state. Those measurable variables constitute an n-dimensional subspace, denoted by $\mathbf{x}$ and the complementary space is thus $N \times D - n$-dimensional marked with variables $\mathbf{y}$. Of course, we may write $\mathbf{X}^t = (\mathbf{x}^t, \mathbf{y}^t)$ if $\mathbf{x}, \mathbf{y}$ are properly arranged. Their dynamics is governed by equation (1), which with a slight abuse of notations could be written in a more convenient form

$$\dot{\mathbf{x}} = \mathbf{F}_o(\mathbf{x}, \mathbf{y}, \mathbf{p}), \dot{\mathbf{y}} = \mathbf{F}_h(\mathbf{x}, \mathbf{y}, \mathbf{p}), \tag{2}$$

where the subscripts $o = (1,2,\ldots,n)^t$ and $h = (n+1, n+2, \ldots, N*D-n)^t$ denote different dimensions of the phase space.

To reconstruct the system, an error function is proposed as an integrated squared difference between the left- and the right- handside of equation (2)

$$\Delta = \int_0^t (\dot{\mathbf{x}} - \mathbf{F}_o(\mathbf{x}, \mathbf{y}, \mathbf{p}))^2 + \beta(\dot{\mathbf{y}} - \mathbf{F}_h(\mathbf{x}, \mathbf{y}, \mathbf{p}))^2 ds \tag{3}$$

in a time interval $[0, t]$, where β is an adjustable hyper-parameter. With correct values of the parameters and orbit points, obviously Δ takes its minimum value 0, indicating the existence of an absolute minimum. While the velocity $\dot{\mathbf{x}}$ could be computed from the known time series of $\mathbf{x}(t)$ [43], the state variables $\mathbf{y}$ as well as their time variations $\dot{\mathbf{y}}$ have to be deduced. The strategy is to systematically modify the parameters $\mathbf{p}$ and the initial conditions of $\mathbf{y}$ so as to minimize the cost function equation (3), treating $\mathbf{x}$ and $\dot{\mathbf{x}}$ as known.

The integration time t should be long enough to exhibit characteristics of the orbit and to suppress small noise in the data. Nevertheless, if t is really large, due to the exponential amplification of deviations in a chaotic system, the error function Δ will become extremely sensitive to the parameters as well as to the initial conditions. To balance the computation, an exponentially decaying factor $e^{-\alpha(t-s)}$ is supplied to the integrand. With this modification, the error function essentially depends only on a time interval ending at t and with a length characterized by the hyper-parameter α. With this decaying factor, the memory of initial

conditions becomes vague with increasing t and we have to employ an alternative way to push the state point to the correct trajectory. For this, we modify the equation of motion of $\mathbf{y}$ in the following manner

$$\dot{\mathbf{y}} = \mathbf{F}_h(\mathbf{x}, \mathbf{y}, \mathbf{p}) + \mathbf{e}, \tag{4}$$

where $\mathbf{e}$ are artificial parameters to gradually correct errors due to the wrong initial conditions, which has the same dimension as $\mathbf{y}$. Later on, as will be seen, the parameters $\mathbf{e}$ are also evolving slowly, sending $\mathbf{y}$ to the right track. Toward the end of the course, $\mathbf{e}$ gradually shrinks to 0, recovering the original equations of motion equation (2).

With all the above consideration, the error functions could be modified to

$$\Delta = \int_0^t e^{-\alpha(t-s)} \left[(\dot{\mathbf{x}} - \mathbf{F}_o(\mathbf{x}, \mathbf{y}, \mathbf{p}))^2 + \beta \mathbf{e}^2 \right] ds. \tag{5}$$

Thus, the problem converts to parameter learning which minimizes the function equation (5). Nevertheless, it is not as simple as it seems. The state variable $\mathbf{y}$ is in fact an implicit function of $\mathbf{p}$ and $\mathbf{e}$ since it is computed from the evolution equation (4). Thus, we have

$$\frac{\partial \Delta}{\partial \mathbf{p}} = \int_0^t e^{-\alpha(t-s)} \sum_{i=1}^n -2(\dot{x}_i - F_i(\mathbf{x}, \mathbf{y}, \mathbf{p})) \left(\frac{\partial F_i(\mathbf{x}, \mathbf{y}, \mathbf{p})}{\partial \mathbf{y}} \frac{\partial \mathbf{y}}{\partial \mathbf{p}} + \frac{\partial F_i(\mathbf{x}, \mathbf{y}, \mathbf{p})}{\partial \mathbf{p}} \right) ds \equiv \mathbf{H}(t)$$

$$\frac{\partial \Delta}{\partial \mathbf{e}} = \int_0^t e^{-\alpha(t-s)} \left(\sum_{i=1}^n -2(\dot{x}_i - F_i(\mathbf{x}, \mathbf{y}, \mathbf{p})) \left(\frac{\partial F_i(\mathbf{x}, \mathbf{y}, \mathbf{p})}{\partial \mathbf{y}} \frac{\partial \mathbf{y}}{\partial \mathbf{e}} \right) + 2\beta e \right) ds \equiv \mathbf{E}(t). \tag{6}$$

From equation (4), we may get the evolution of the parameter dependence

$$\frac{d}{dt} \left(\frac{\partial \mathbf{y}}{\partial \mathbf{p}} \right) = \frac{\partial \mathbf{F_h}}{\partial \mathbf{y}} \frac{\partial \mathbf{y}}{\partial \mathbf{p}} + \frac{\partial \mathbf{F_h}}{\partial \mathbf{p}} - \alpha \frac{\partial \mathbf{y}}{\partial \mathbf{p}}$$

$$\frac{d}{dt} \left(\frac{\partial \mathbf{y}}{\partial \mathbf{e}} \right) = \frac{\partial \mathbf{F_h}}{\partial \mathbf{y}} \frac{\partial \mathbf{y}}{\partial \mathbf{e}} + 1 - \alpha \frac{\partial \mathbf{y}}{\partial \mathbf{e}}. \tag{7}$$

In the above two equations, the last terms that contain α are supplied to damp the parameter dependence of earlier times. With the increase of integration time, this dependence could become extremely sensitive in a chaotic motion if $\alpha = 0$. Therefore, those two terms will keep a finite memory as done in equation (5).

The evolution of parameters can be chosen along the gradient direction, thus

$$\frac{d\mathbf{p}}{dt} = -2\gamma \mathbf{H}, \frac{d\mathbf{e}}{dt} = -2\gamma \mathbf{E} \tag{8}$$

where γ is the learning rate and the variables $\mathbf{H}, \mathbf{E}$ are defined in equation (4) and satisfy the equations

$$\frac{d\mathbf{H}}{dt} = -\alpha \mathbf{H} + \mathbf{H}_1, \frac{d\mathbf{E}}{dt} = -\alpha \mathbf{E} + \mathbf{E}_1, \tag{9}$$

where $\mathbf{H}_1, \mathbf{E}_1$ is the integrand apart from the decay factor in the expression $\mathbf{H}, \mathbf{E}$ displayed in equation (6).

With the above set of evolution equations for $\mathbf{y}, \frac{\partial \mathbf{y}}{\partial \mathbf{p}}, \frac{\partial \mathbf{y}}{\partial \mathbf{e}}, \mathbf{p}, \mathbf{e}, \mathbf{H}, \mathbf{E}$ and the given time series of $\mathbf{x}$, we may drive the unknown variables to recover both the parameters and the orbit (see figure 1). The initial conditions of these equations are given somewhat arbitrarily although prior information could be used to speed up the convergence. Assuming the number of unknown parameters $\mathbf{p}$ is n_p, then the total number of parameters is $n_c = n_p + n_e = n_p + N * D - n$ where $n_e = N * D - n$ is the number of the auxiliary parameters $\mathbf{e}$. Therefore, the total number of equations to solve is $n_t = (N * D - n)(1 + n_c) + 2n_c$, since we care about each hidden variable as well as their derivatives with respect to all the parameters and the evolution of each parameter needs two equations (check equations (8) and (9)). It is easy to see that the number of equations is quadratic in the number of hidden variables $\mathbf{y}$ but linear in unknown parameters $\mathbf{p}$. Hence, the more information we have, the less effort is needed to solve these equations, being consistent with the common intuition. Thus, the complexity increases quadratically with the number of hidden variables, which makes this type of problems very hard.

2.2. Numerical simulations

In the following, several well-known examples are used to illustrate application of the above formulation for recovery of system dynamics and parameters. The absolute error of an estimated parameter is calculated as $|parm - parm'|$, where $parm$ is the inferred result and $parm'$ is the true value. First, as usual we try the Lorenz system and check how the performance of the new algorithm depends on various hyper-parameters.
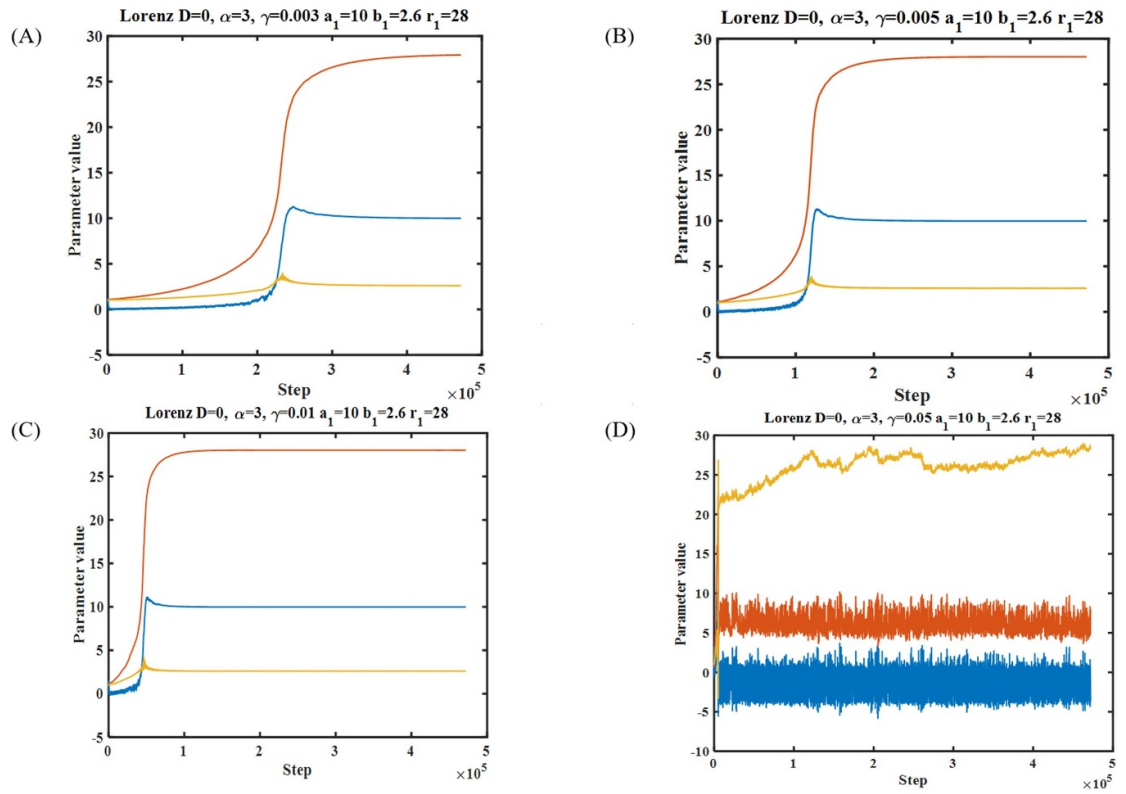
Figure 2. Reconstruction with different γ at $t_{\text{unit}} = 0.01, \alpha = 3$, and $D = 0$. (A) $\gamma = 0.003$. (B) $\gamma = 0.005$. (C) $\gamma = 0.01$. (D) $\gamma = 0.05$.

2.2.1. Lorenz system

The Lorenz system is a famous deterministic dissipative system [61], which is written as

$$
\begin{aligned}
\dot{x} &= a(y - x), \\
\dot{y} &= rx - y - xz, \\
\dot{z} &= xy - bz.
\end{aligned}
\tag{10}
$$

With the often-chosen parameters $a = 10, b = 2.6, r = 28$, this system displays very typical chaotic behaviour with the maximum Lyapunov exponent close to 1. If the time series of x, y, z are all known, it is not hard to recover the coefficients of the polynomials on the right hand side of equation (10) [20, 22, 35, 42, 43]. Instead, if only one variable, say x, is observed, the recovery becomes much harder [59]. Here, the new formulation is applied with two hidden variables y, z, three unknown parameters a, r, b and two auxilliary parameters e_2, e_3. Thus, according to the above discussion, the total number of equations is $n_t = 2 \times (1 + 3 + 2) + (3 + 2) \times 2 = 22$. In the following simulations, $[a, r, b] = [1, 1, 1], [y, z] = [1, 1]$ are conveniently used as the initial conditions for the corresponding variables. With the above algorithm, the three unknown parameters a, r, b converge well as shown in figures 2(A)–(C)and 3(C), (D). For different hyper-parameters, the convergence seems to follow similar patterns. Initially, the unknown parameters change steadily but at some point start to rush to the true values and then remain there throughout the simulation. Nevertheless, the convergence is different in disparate conditions and could even go awry if improper hyper-parameters are chosen as displayed in figures 2(D) and 3(A), (B).

In this example, the integration time step is $t_{\text{unit}} = 0.01$. As can be seen from the figures, a total number of $N_{step} \sim 10^5$ steps are needed for the convergence, which is determined by the two hyper-parameters α, γ. Currently, we take $\alpha = 3$, indicating a memory of $1/\alpha = 0.33$ which is about a third of the Lyapunov time of the system. The learning rate is chosen to be $\gamma = 0.01$, about one hundredth of the Lyapunov exponent. γ is used to control the update of the parameters in the evolution equation and should keep small compared with the time rate of the system state (y, z) in order not to disrupt the generic dynamics of the original system. As shown in figures 2(A)–(C), a reduction of γ extends the learning time and requires more data for convergence. Nevertheless, if the available time series has a limited length, a technique of data recycling may be employed for a successful reconstruction as shown in appendix 'Recycling of data'. On the other hand, if γ
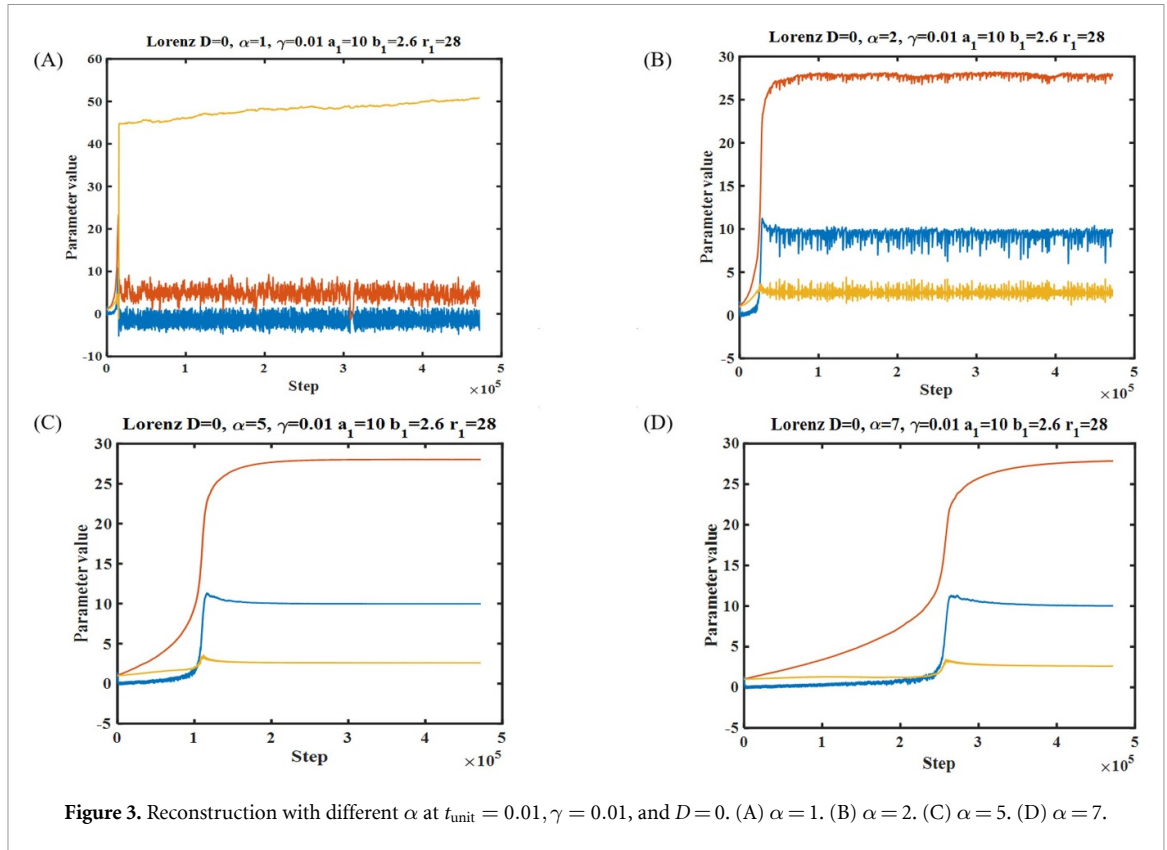
Figure 3. Reconstruction with different α at $t_{\text{unit}} = 0.01$, $\gamma = 0.01$, and $D = 0$. (A) $\alpha = 1$. (B) $\alpha = 2$. (C) $\alpha = 5$. (D) $\alpha = 7$.

is overly large, the parameters change rapidly and the original dynamics may be altered so that no convergence results, as displayed in figure 2(D). So during the reconstruction, the learning rate should be chosen properly to ensure both the stability and efficiency.

The decay rate α should not be too small in order to bound the integrals in equation (6). On the other hand, if α is too large, the memory becomes really short and these integrals highly depend on the current state which varies rapidly and easily overruns the slow parameter changes. In figure 3, the influence of the decay rate α on the convergence is displayed. As α increases, the learning becomes slow and more data is needed, just like reducing γ. When α is relatively small (figure 3(A)), the effective integration time becomes long and the integral will be very sensitive to parameter changes, causing instability in the system, unless the learning rate γ is really small and averaging out the fast oscillations. Therefore, for a system with a large Lyapunov exponent, when the learning rate is fixed, it is necessary to apply a proper decay rate to reduce the sensitivity of the parameters (appendix figures 21 and 22). In general, it is important to adjust the learning and the decay rate to keep a balance between the two so as to achieve a stable and fast convergence.

From the above discussions, it seems that the current reconstruction scheme is quite fragile. Actually, it is not and able to work quite well even in the presence of noise. Next, we investigate the impact of noise on the reconstruction since it inevitably exists in every experiment or observation. For this purpose, we add a noise term to each of the evolution equation (10), which describes white noise with the following characteristics

$$< \eta_i(t) > = 0, \tag{11a}$$

$$< \eta_i(t)\eta_j(t') > = 2D\delta_{ij}(t - t'), \tag{11b}$$

where $\langle \cdot \rangle$ denotes an ensemble average and the delta function is a Kronecker one for the discrete index i, j, and a Dirac one for the continuous time variables $t - t'$. D is the noise intensity and when $D = 0$ we recover the deterministic autonomous system.

Even though a noisy orbit is generated with equation (10) with noise ($D \neq 0$), we still use the same deterministic formulation in the previous section. As an example, suppose that only the noisy $x-$orbit is given and we still aim to deduce the unknown parameters and the hidden variables $y(t), z(t)$. If the noise is not very strong, $x(t)$ may be directly substituted into the 22 reconstruction evolution equations mentioned above, together with the $\dot{x}(t)$ computed with any fair smoothening technique [43]. As a result, only the
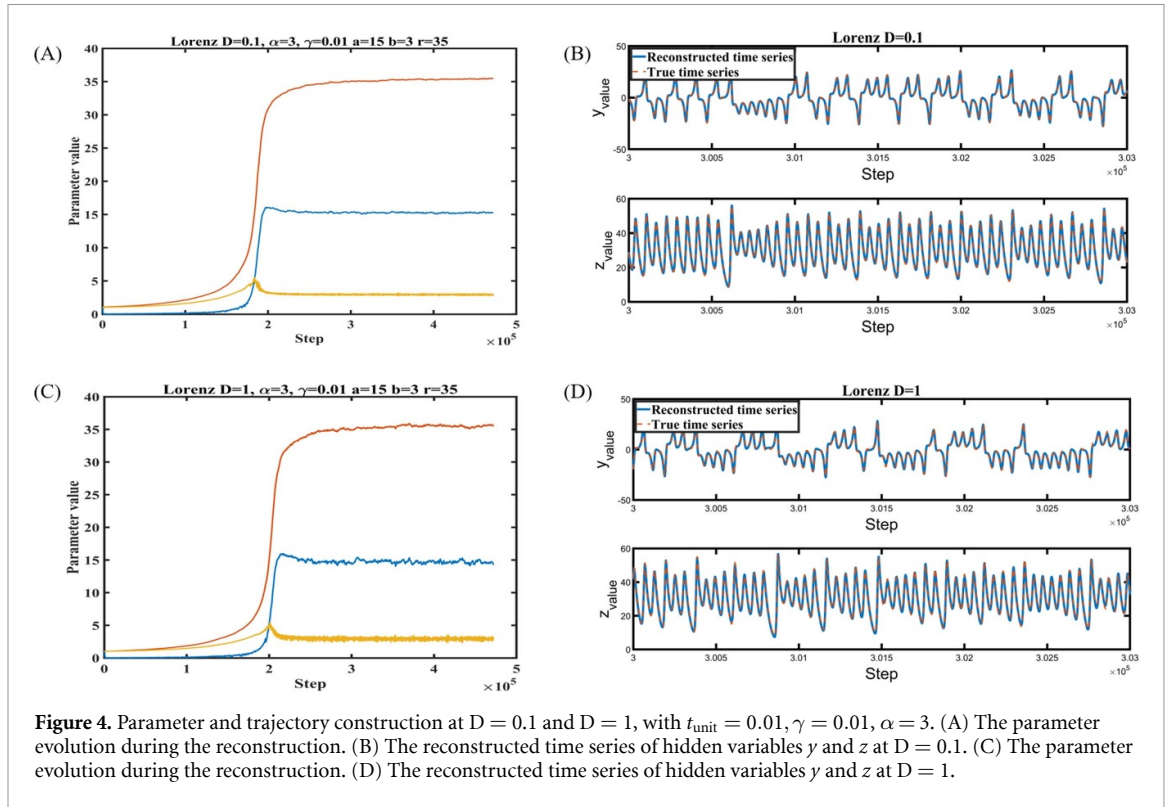
Figure 4. Parameter and trajectory construction at D = 0.1 and D = 1, with $t_{unit} = 0.01$, $\gamma = 0.01$, $\alpha = 3$. (A) The parameter evolution during the reconstruction. (B) The reconstructed time series of hidden variables y and z at D = 0.1. (C) The parameter evolution during the reconstruction. (D) The reconstructed time series of hidden variables y and z at D = 1.

Table 1. The dependence of parameter reconstruction on noise in the Lorenz system.

	$D=0$	$D=0.1$	$D=1$	$D=5$
$r=35$	$r=35.19$	$r=35.35$	$r=35.52$	$r=34.56$
$b=3$	$b=2.95$	$b=2.94$	$b=2.93$	$b=3.07$
$a=15$	$a=15.26$	$a=15.26$	$a=14.68$	$a=13.55$

average orbit is reconstructed with the unknown parameters. Figure 4 shows the reconstruction results at different noise intensities, i.e. $D = 0.1$ in figures 4(A) and (B) and $D = 1$ in (C) and (D). The convergence pattern of the parameters displayed in figures 4(A) and (C) looks very similar to that in the deterministic case but small jittering exists due to noise perturbation, which is more clearly seen in (C) for $D = 1$. After the small oscillations are averaged out, we get a fairly good estimate of the parameters, which is listed in table 1. In fact, the results of extra tries are also displayed in the table up to $D = 5$. In general, the estimation gets worse for greater noise intensity as expected but the errors do not appear monotone. The reconstructed orbits $y(t), z(t)$ are displayed in figures 4(B) and (D), which seem almost overlapping with the true orbits, just as in the deterministic case.
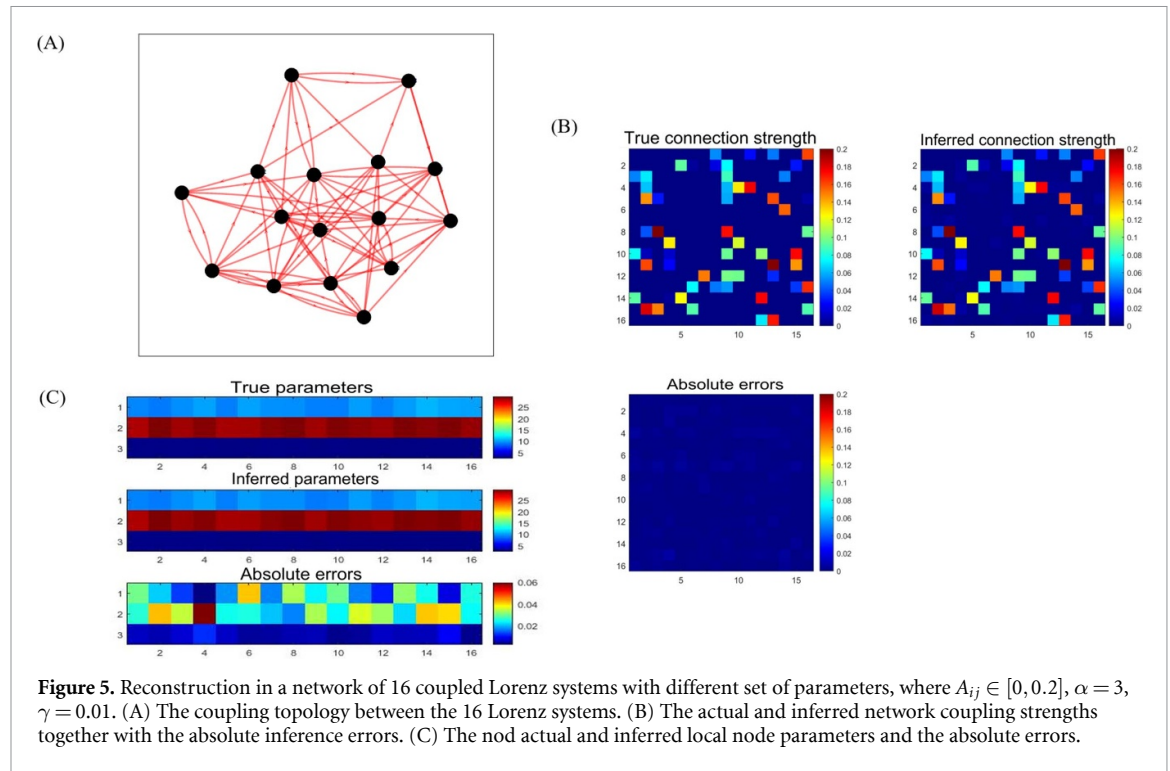
From the above analysis, it can be seen that the new construction scheme is quite robust against noise as long as it is not too strong. Next, the scheme will be applied to several typical examples in nonlinear dynamics with high dimensions to show its validity.

2.2.2. The N coupled Lorenz oscillators
First, we consider the inference problem in a network with N nodes each of which is occupied by a Lorenz system. The connection between nodes marks their interaction. More specifically, the ith node is governed by the equation below

$$\dot{x}_i = a_i(y_i - x_i) + \eta_i(t),$$

$$\dot{y}_i = r_i x_i - y_i - x_i z_i + \sum_{j=1}^{N} A_{ij}(y_j - y_i) + \eta_i(t),$$

$$\dot{z}_i = x_i y_i - b_i z_i + \eta_i(t) \tag{12}$$

where $\eta_i(t)$ is the above-mentioned Gaussian white noise and A_{ij} is the interaction matrix. Here, without less of generality, the interaction is assumed to take place between $y-$components. The local parameters are

Figure 5. Reconstruction in a network of 16 coupled Lorenz systems with different set of parameters, where $A_{ij} \in [0, 0.2]$, $\alpha = 3$, $\gamma = 0.01$. (A) The coupling topology between the 16 Lorenz systems. (B) The actual and inferred network coupling strengths together with the absolute inference errors. (C) The nod actual and inferred local node parameters and the absolute errors.

randomly selected from intervals on the positive real axis: $a_i \in [10, 12], r_i \in [28, 30], b_i \in [2.6, 2.8]$. The noise intensity $D = 0.01$ and the sampling interval $t_{\text{unit}} = 0.01$. Suppose that only the $x-$components are measured. Our goal is to reconstruct the coupling matrix A_{ij} and all the local parameters $\{a_i, b_i, r_i\}_{i=1,2,...,N}$ as well as the time-dependent hidden variables.

With the above scheme, for an Erdős-Rényi (ER) network with $N = 16$, the coupling matrix and the node parameters are correctly deduced as displayed in figure 5. For simplicity, we take the initial conditions of interaction matrix A_{ij} to be zero unless otherwise stated, and the initial values of local parameters are $[a_i, r_i, b_i] = [1, 1, 1], [y_i, z_i] = [1, 1]$. After the reconstruction, the absolute errors of the computation in the coupling matrix are of the order of 10^{-3}, and around 10^{-2} in the node parameters. The time series of all hidden variables are also reconstructed at the same time, which is not shown here for brevity. Please check figure 17 in the appendix for one example.

When N is small, we may start with a full matrix $[A_{i,j}]$ containing N^2 elements to be determined. The network structure can be drawn from the non-zero elements of $[A_{i,j}]$ after the reconstruction and we do not need any prior information about the network topology. Nevertheless, this type of analysis is not scalable since the quadratic increase of the number of the unknown matrix elements with the network size prevents their efficient estimation based only on the linearly increasing information. In fact, the current brute-force method works up to $N \sim 20$. For larger networks, extra information about the network topology helps a lot, especially those that cut down the number of unknown matrix elements. For example, if the structure of the network is approximately known (which may still contain some spurious links), our computation remains valid as long as the number of links increases not that fast.

Here, we tried the reconstruction on a small-world network with $N = 100$ and an average degree $K = 4$ as portrayed in figure 6(A). For an easy comparison, the non-zero coupling is allowed taking only four values $0.1, 0.3, 0.5, 0.8$. The results are displayed in figures 6(B) and (C) which agree well with the original values. The convergence of the local node parameters looks quite similar as in precious cases (see figure 6(B)). However, the evolution of the couplings starts more erratically. The fluctuations grow larger and larger and culminate just before a sudden convergence to the true values (see figure 6(C)). The whole process closely resembles what happens in chaos control: only when the system state gets close to the stable manifold of the true dynamics and the true parameters does the gradient descent scheme in equation (8) start to work and drive the wandering point to the right solution [62].

In the current example, we have 700 unknown system parameters, including 300 node parameters and 400 coupling strengths. In addition, 200 auxiliary parameters **e** are needed to drive $y_i(t), z_i(t)$ to the right trajectories. Altogether, we have 900 hundred parameters, which will generate over 180 000 equations in the new formulation, if we exactly follow the procedures in the theory part. However, a convenient approximation could be adopted based on the given network structure and the observation that the influence
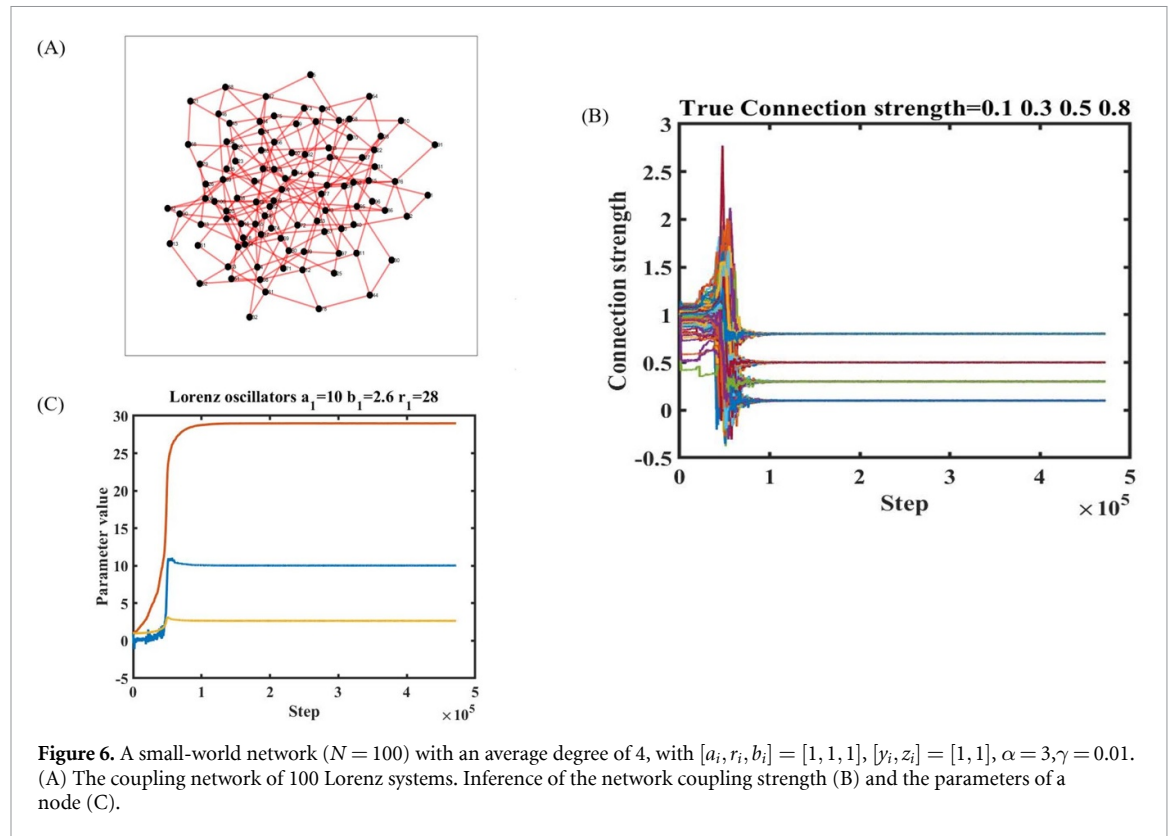
Figure 6. A small-world network ($N = 100$) with an average degree of 4, with $[a_i, r_i, b_i] = [1, 1, 1]$, $[y_i, z_i] = [1, 1]$, $\alpha = 3, \gamma = 0.01$. (A) The coupling network of 100 Lorenz systems. Inference of the network coupling strength (B) and the parameters of a node (C).

of a node on other nodes decays quickly in an asynchronous state. According to the equation of motion, each of the above-mentioned parameters could be assigned to a node or a connection between nodes. To the lowest order, it is reasonable to assume that the dynamics at a node mainly depends on the parameters of neighbouring nodes or edges. With this assumption, the partial derivatives with parameters are taken to be zero if they locate outside the immediate neighbourhood. As a result, the number of equations is greatly cut down, which in fact grows only linearly with the size of the network with a fixed average degree.

As the true values of the parameters or state variables are obtained through a co-evolution with the observed time series, the proposed method can be applied online to monitor changes in network topology or node dynamics if the change does not happen so frequently. As an example, the above formulation is directly applied to the network with 8 nodes displayed in figure 17(A) in the appendix. Starting from a somewhat arbitrary initial guess value 1, the coupling converges to the true value 0.1 after a quite long transient. Later on, as the coupling is switched to 0.5 in the network, the reconstruction algorithm detects and quickly follows this change. It is interesting to see that the reaction is really fast this time and the transient is almost unnoticeable as shown in figure 7(A). The node parameters that change at the same time can be similarly observed, as shown in figure 7(B). One reason for the quick response for the parameter change is that the state variables $\{y_i(t), z_i(t)\}_{i=1,2,\ldots,N}$ remain close to the true values in this process so that the convergence seems monotone and exponential!

2.2.3. The FitzHugh–Nagumo networks

A salient feature of the FitzHugh–Nagumo (FHN) system is the excitability—once the potential exceeds a threshold, the system state immediately shoots to a very high value and then relaxes back to the base. Therefore, the temporal pattern is quite different from that of a Lorentz system. Is it still possible to apply our method to this new situation? Below, we check the continued validity of the current scheme.

The FHN equation is a second-order nonlinear system used to simulate the dynamics of a neuron. To model neural networks, these neurons connect to each other and interact with their neighbours. The equations of motion can be written as

$$\dot{v}_i = v_i - v_i^3 - w_i - \sum_{j=1}^{N} A_{ij}(v_j - v_i) + \Gamma_{1,i}$$

$$\dot{w}_i = a_i + b_i + c_i w_i + \Gamma_{2,i}$$

(13)
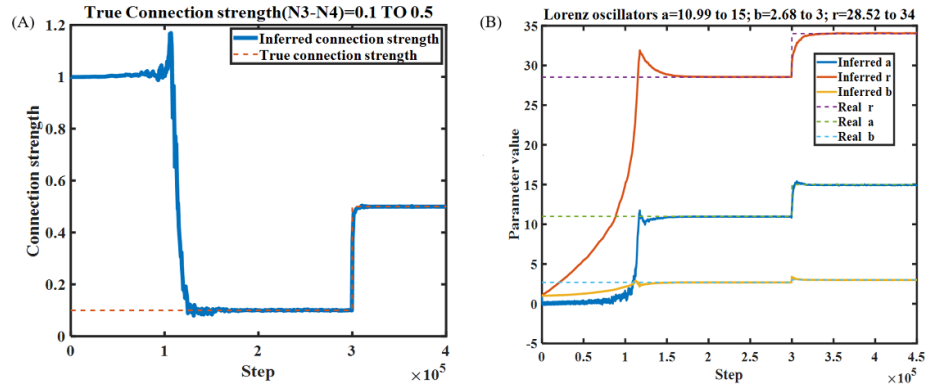
Figure 7. Monitoring changes in network topology and node dynamics. (A) The change in coupling strength between nodes over time. (B) The change of parameters in node dynamics over time.
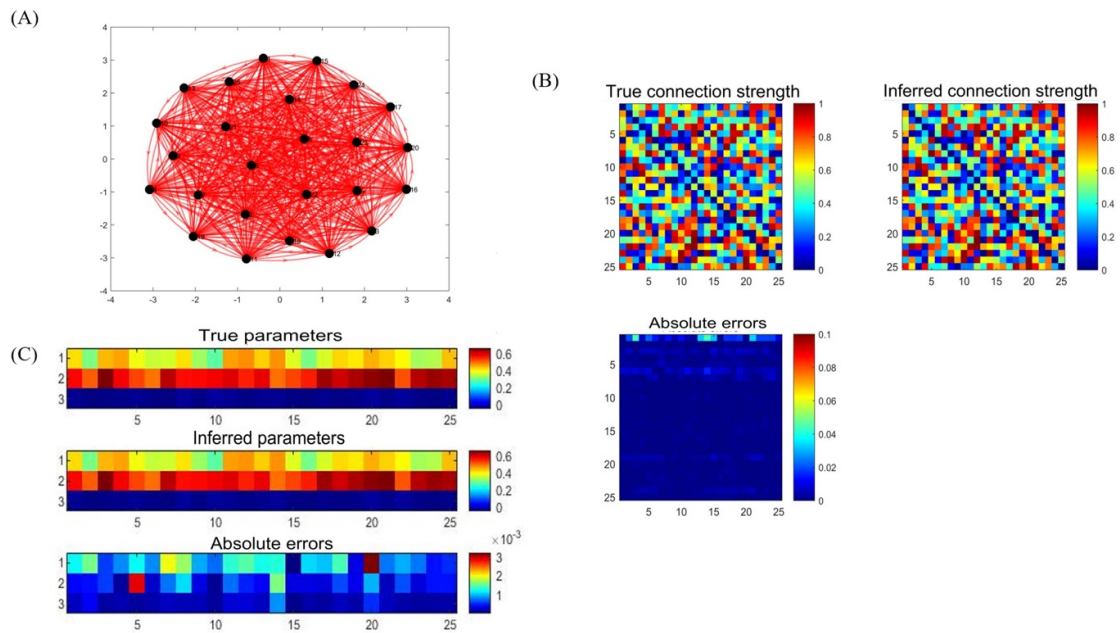

Figure 8. Reconstruction in a network of 25 FHN nodes with $\gamma = 0.5$, $\alpha = 5$. (A) The network topology. (B) The actual and inferred network coupling strengths and the absolute errors. (C) The actual and inferred local parameters and the absolute errors.

where $i = 1, 2, \ldots, N$ denotes different neurons, v_i is the rate of change of the membrane potential, w_i is the slowly changing recovery variable of the ith neuron and N is the number of neurons. In this model, the dynamical behaviour of the system largely hinges on the coefficients
$a_i \in [0.3, 0.5], b_i \in [0.5, 0.7], c_i \in [-0.04, -0.03]$ and the interaction weight $A_{ij} \in [0, 1]$. $\Gamma_{1,i}, \Gamma_{2,i}$ are noise terms of the two components, which are usually taken to be white satisfying equation (11), with $D = 0.01$.

In practice, only the potential $v_i(t)$ of each neuron is subject to a direct measurement. Is it possible to recover $w_i(t)$ and all the parameters in the equation with this information? figure 8(A) displays a ER network with $N = 25$ from which the measured data was obtained. As usual, we apply the above scheme by initially setting all unknown coupling strength A_{ij} to 1 and the local system parameters to 0. With this network size, all the parameters could be obtained directly without other assumptions. The obtained coupling strengths are depicted in figure 8(B) with the absolute error of the order of 10^{-2}, while the local parameters are portrayed in figure 8(C) with the absolute error of the order of 10^{-3}. Figure 9(A) shows the convergence of local parameters for a node from the initial values 0 to the correct values, where the oscillatory character at the beginning is still clearly seen. Figure 9(B) displays the time series of $v_i(t)$ where the profile is contaminated partly by visible noise. However, the reconstructed time series of w_i seems smooth and overlaps with the true one almost perfectly.

For small-world FHN networks with known structures, for example, the one in figure 6(A), how does the reconstruction accuracy depend on the network size? If the average degree is fixed to 4, this dependence is
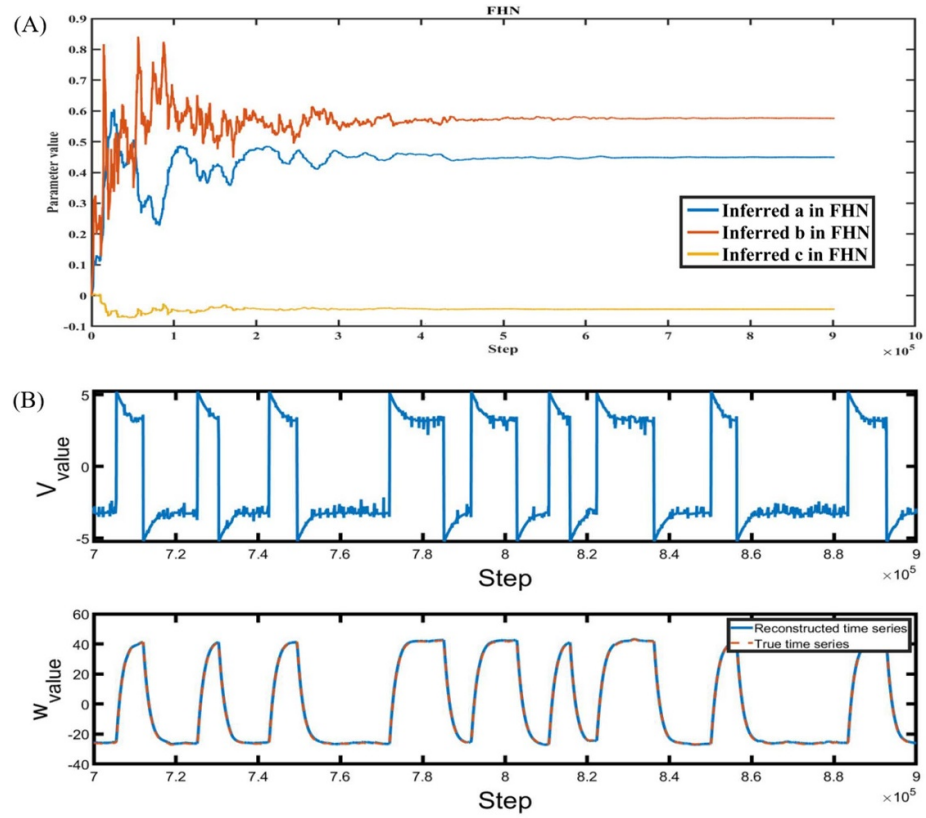
Figure 9. The parameter reconstruction and the evolution of state variables at one node of the FHN network. (A) The parameter reconstruction. (B) The measured time series of *v* and *w* together with the reconstructed *w*.

shown in figure 10(A). It is easy to see that the inference accuracy can be maintained above 95% after averaging over 20 realizations. The accuracy is defined as [20]

$$Ac_\rho = \frac{1}{K \times N} \sum_{i,j} H\left((1-\rho) - \Delta A_{ij}\right), \tag{14}$$

where ρ is the required accuracy and we take $\rho = 0.1$ here. H is the Heaviside function and the relative error between the reconstructed ($\hat{A}_{ij}$) and the real coupling (A_{ij}) is:

$$\Delta A_{ij} = \left|\hat{A}_{ij} - A_{ij}\right| / \left(A_{ij}\right). \tag{15}$$

The main source of error is the synchronization among part of the nodes, which makes it hard to distinguish them and thus leads to incorrect inferences. More discussion will be made below. On the other hand, if we keep the network size and increase the average degree, the inference will become harder since more couplings need deducing. Nevertheless, according to figure 10(B), the accuracy still maintains near the value 95% throughout.

2.2.4. The Kuramoto oscillators

The ability to deduce a large number of parameters from the time series may be applied to the reconstruction of networks with hidden nodes, about which little information is known. Even their existence has to be deduced from the measurement on neighbouring nodes as done in previous works to determine the possible connections of hidden nodes to the known ones in the network [51–54]. Nevertheless, this type of deduction is able to confirm the existence of hidden nodes but the precise connections may not be inferred very accurately. In our scheme, all possible links may be included in the deduction as long as the node's existence is expected. The spurious ones will give a weight near zero and the weights will be recovered for the true ones. Also, the dynamics and internal parameters of the hidden nodes could be deduced as well.

Here, Kuramoto oscillator model [63] defined on networks will be used as an example for this computation. Each oscillator in the model is described with a phase, so $D = 1$. The coupling between them
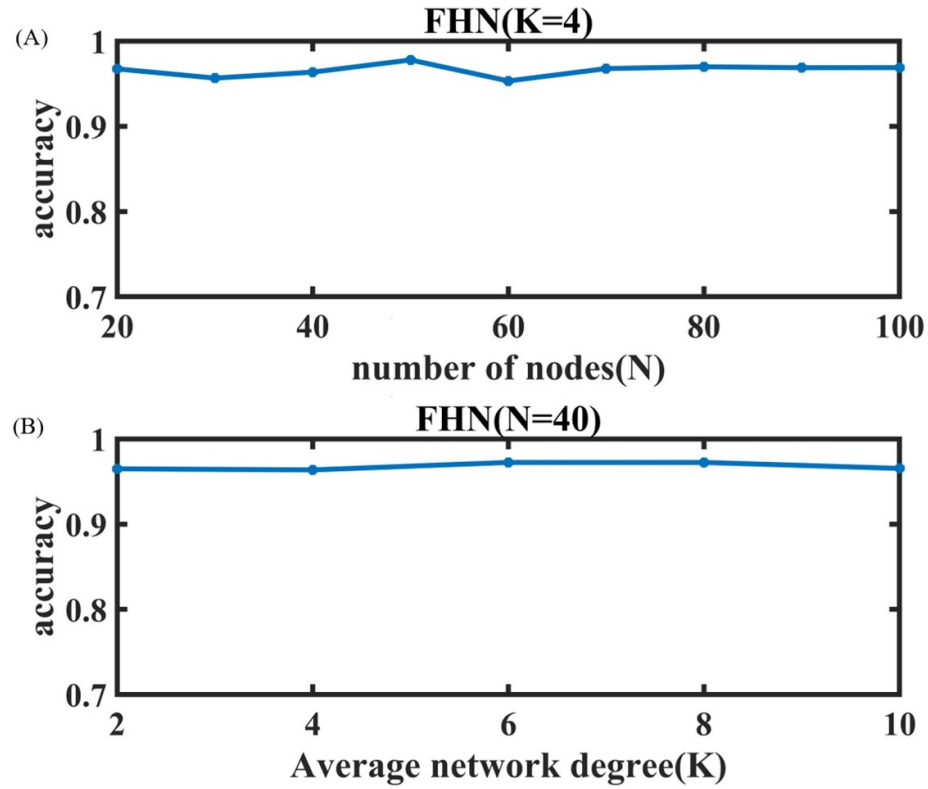
Figure 10. How the accuracy of network reconstruction depends on (A) the network size and (B) the average degree.

induces synchronization at various levels, which has been studied comprehensively [64, 65]. Consider a system with N oscillators

$$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j=1}^{N} a_{ij}\sin\left(\theta_j - \theta_i\right), (i = 1, \ldots, N),\tag{16}$$

where θ_i and ω_i are the phase and the natural frequency of the ith oscillator, $K > 0$ is the coupling strength and a_{ij} is the adjacency matrix of the network. When node i and node j are connected, $a_{ij} = 1$; otherwise $a_{ij} = 0$.
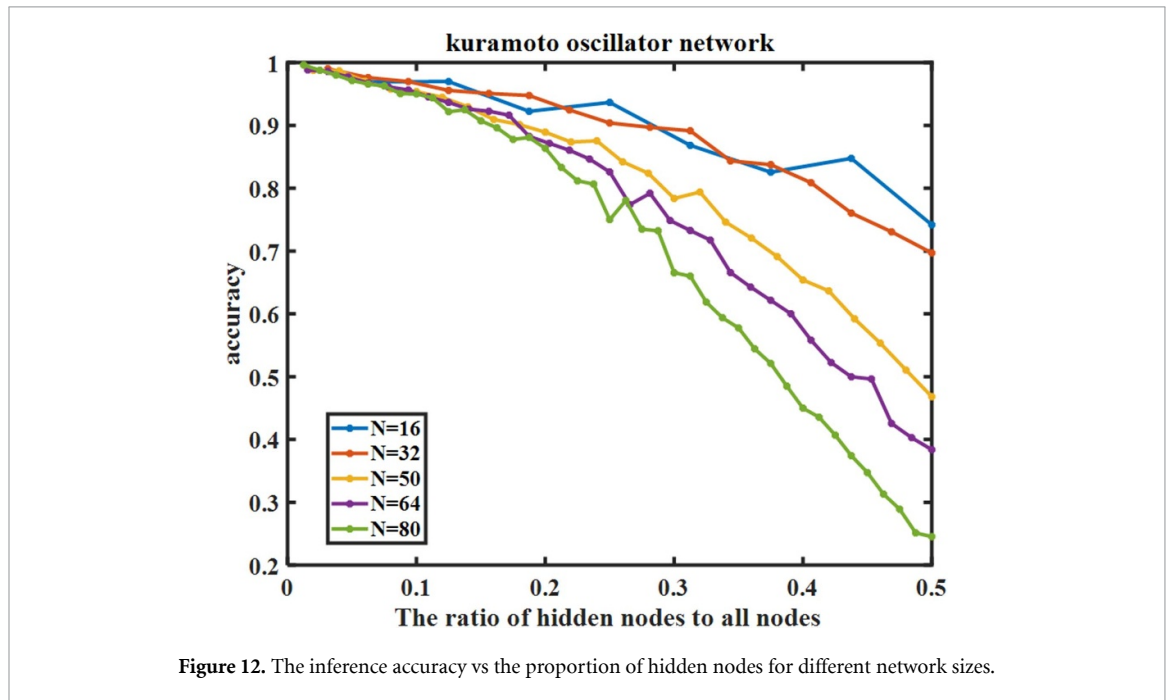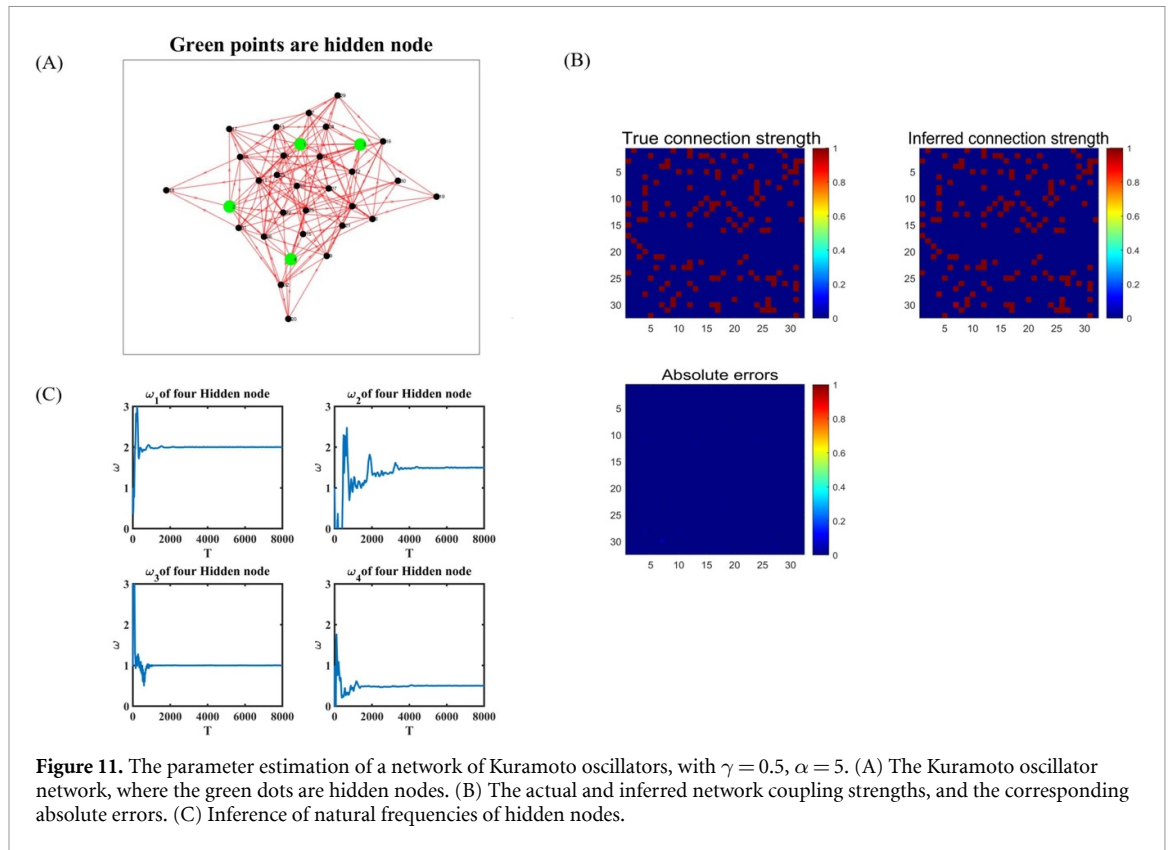
To characterize synchronization, in the literature [65], an order parameter R is defined as follows
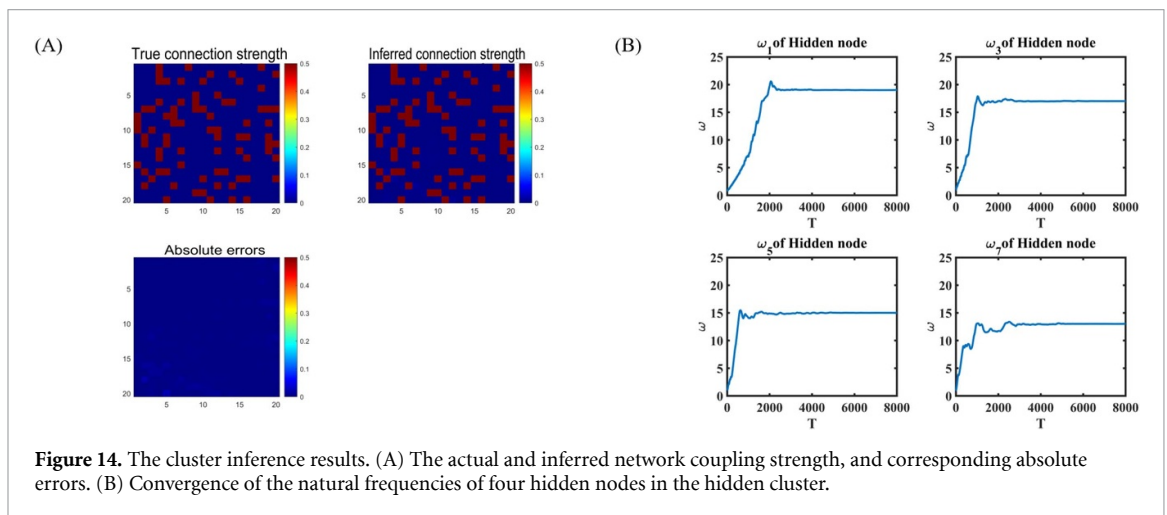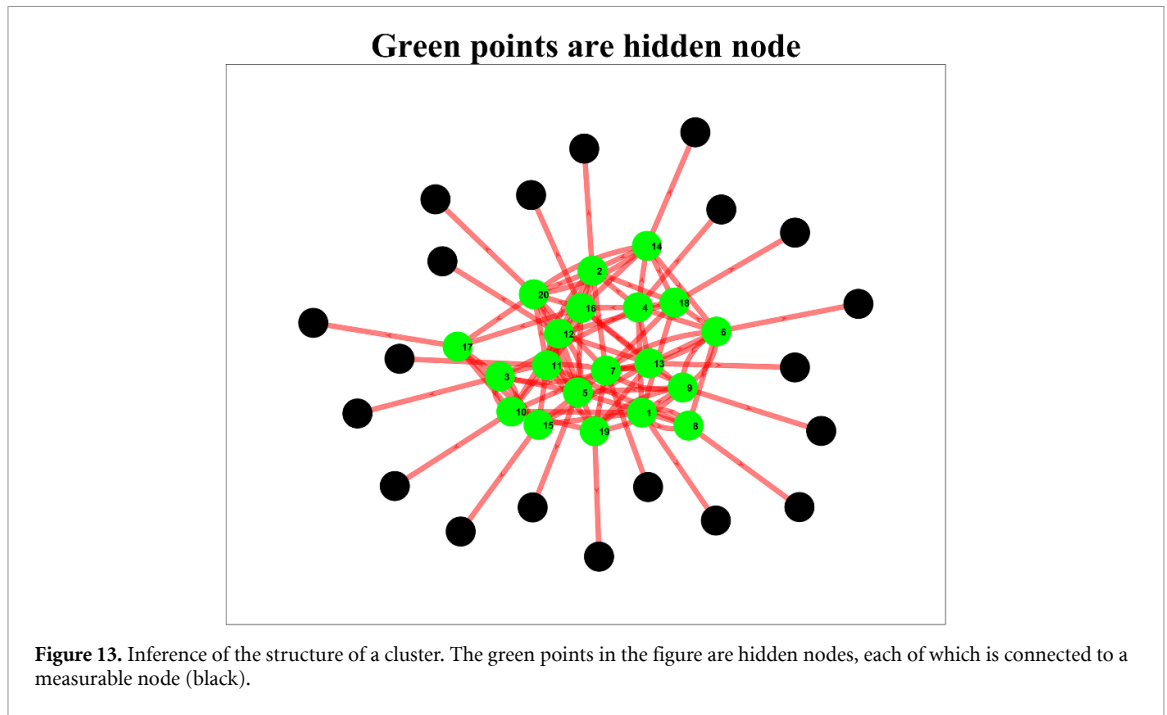
$$Re^{i\psi} = \frac{1}{N}\sum_{j} e^{i\theta_j},\tag{17}$$

where $R, \psi \in \mathbb{R}$. If $K \ll 1$, each oscillator rotates near its natural frequency and $R \sim 0$. With increasing K, the average frequencies of connected oscillators approach the average natural frequency in general and some of them may start to rotate together, reaching a partial synchronization such that $R > 0$. For sufficiently large K, all the oscillators become synchronous and rotate with a common frequency $\sum_i \omega_i/N$, so that $R \sim 1$. During the course of dynamics or parameter deduction, synchronization is harmful as seen previously and will be explained below.

In the current example, 32 oscillators are connected to form an ER network as shown in figure 11(A) and ω_i's are picked up from a uniform distribution in the interval $[0, 20]$. The coupling strength $K = 1$ is not very large to prevent synchronization. We assume that four hidden nodes exist in the network. Their dynamics and natural frequencies are not revealed to us, while the time series of the remaining nodes are known. First, the very existence could be deduced together with their rough connections with existing methods in the literature [51]. Next, the whole network structure and the natural frequencies of these hidden nodes are reconstructed with the formalism introduced in this manuscript. In the following, the hidden nodes are randomly selected from the network with no neighbouring pair.

In this example, as the number of nodes is not large, as long as the very existence of hidden four nodes are known, we may start with a full connection matrix in the reconstruction. That is, the initial values of all

Figure 11. The parameter estimation of a network of Kuramoto oscillators, with $\gamma = 0.5$, $\alpha = 5$. (A) The Kuramoto oscillator network, where the green dots are hidden nodes. (B) The actual and inferred network coupling strengths, and the corresponding absolute errors. (C) Inference of natural frequencies of hidden nodes.

Figure 12. The inference accuracy vs the proportion of hidden nodes for different network sizes.

entries of $[a_{ij}]$ are assumed to be 0.5,and the natural frequency ω_i's start with the value 1. In figure 11(B), a comparison is made between the simulation results and true values of the natural frequencies. The error seems small, not exceeding a few percent. The convergence of the natural frequencies of the hidden nodes is displayed in figure 11(C). The familiar oscillatory profiles appear again but they decay much faster than in previous examples. In figure 12, it is easy to see that the accuracy of the reconstruction (see equation (14)) decreases with the increase of hidden nodes. The larger the network, the faster the decay. Nevertheless, from the figure we see that when the hidden nodes are less than 15%, no matter how large the network size is, the accuracy is higher than 90%.

Figure 13. Inference of the structure of a cluster. The green points in the figure are hidden nodes, each of which is connected to a measurable node (black).

Figure 14. The cluster inference results. (A) The actual and inferred network coupling strength, and corresponding absolute errors. (B) Convergence of the natural frequencies of four hidden nodes in the hidden cluster.

In fact, our method is even able to infer the structure of a cluster of hidden nodes. In figure 13, there are 20 hidden nodes (green dots) forming a small cluster. The structure as well as all the parameters is unknown and none of them is observed directly. We can only measure their neighbouring nodes (black ones). From the measured time series, the natural frequencies of nodes in the hidden cluster and the coupling strengths among nodes can be inferred quite accurately, which is clearly displayed in figure 14(A). A full connection between hidden nodes is assumed at the start and the algorithm will computed the actual strengths of the links: about 0 for the non-existing ones and 0.5 for the true connections. The convergence seems fast (see figure 14(B)).

2.2.5. Synchronization undermines deduction

If the network is fully synchronized, all nodes behave in essentially the same way. The time series look similar apart from a phase or a scaling factor, so it is hard to carry out the inference (see appendix figure 23). Generally speaking, the more synchronized the network, the harder it is to infer the structure or coupling. Taking the Kuramoto oscillator network as an example, the natural frequency ω of each node follows the uniform distribution $W(0, \mu)$, and a larger μ indicates that the dynamics has higher heterogeneity. As shown in the figure 15, for $\mu \in [1, 10]$, with the increase of μ, the inference accuracy increases significantly. When μ reaches 10, the accuracy is over 95%, and fluctuates around this value upon further increment of μ.
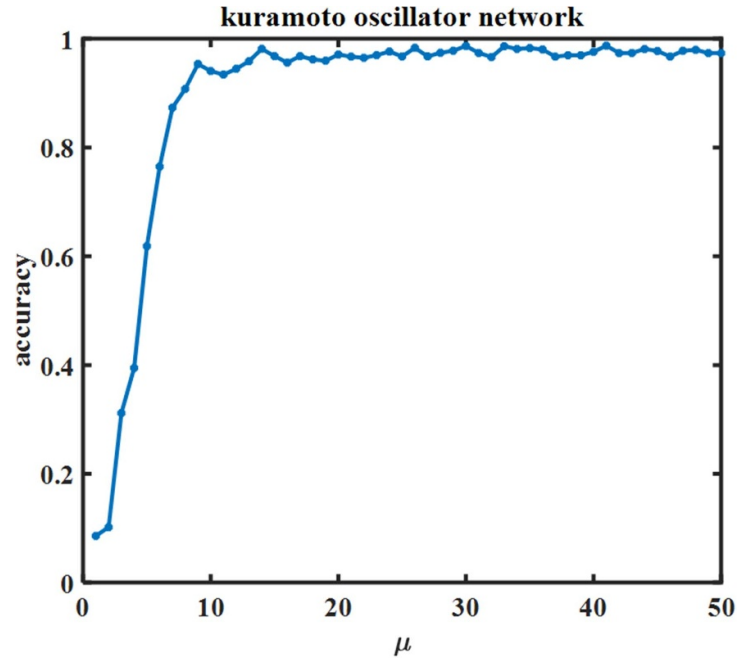
Figure 15. The inference accuracy for the Kuramoto network ($N = 32$) with different natural frequency ranges.
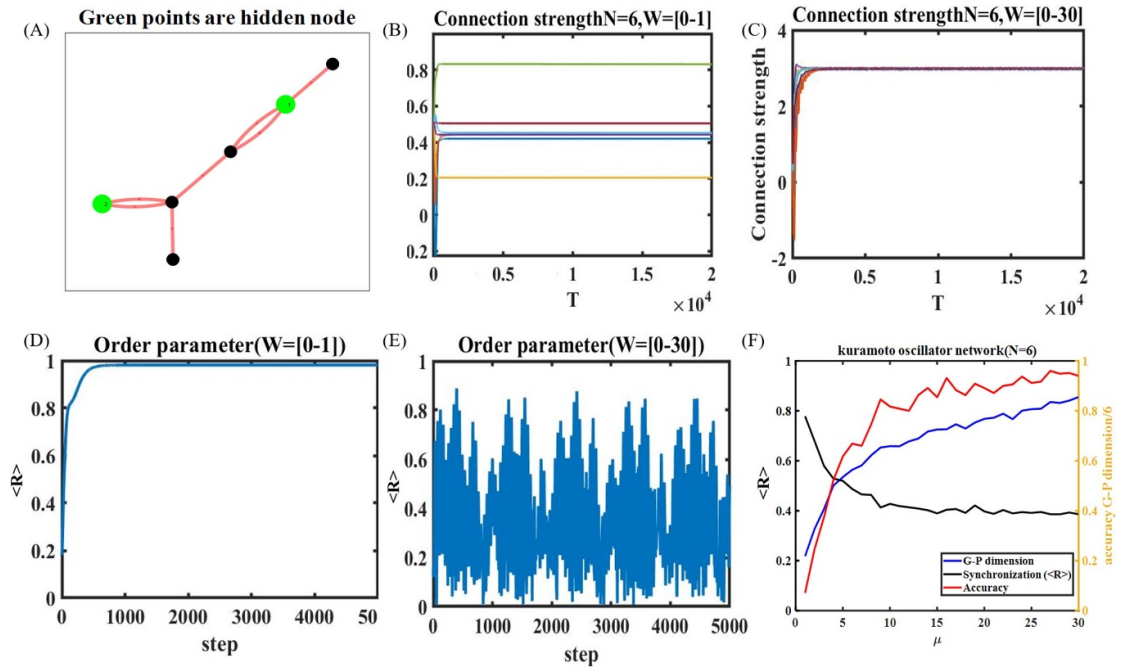
Figure 16. Reconstruction *vs* synchronization. (A) A network of six Kuramoto oscillators with coupling strength $K = 3$, where the green dots are hidden nodes. (B) The inference of the coupling strengths at $\mu = 1$. (C) The inference of the coupling strengths at $\mu = 30$. (D) The convergence of the order parameter $< R > \rightarrow 1$ at $\mu = 1$. (E) The oscillatory behaviour of the order parameter at $\mu = 30$. (F) The dependence on μ of the G-P dimension (normalization by dividing with the maximum value 6), the order parameter $< R >$ and the reconstruction accuracy.

In order to address in more detail the effect of synchronization on inference, we study a network of six oscillators coupled in a specific configuration as shown in figure 16(A), two of which are assumed to be hidden nodes. The natural frequencies of the oscillators conform to a uniform distribution in $[0, 1]$ and the coupling strength is set to 3. In figure 16(B), the coupling strengths indeed converge, but to completely wrong values, since in this case the oscillators fully synchronize. Nevertheless, if the distribution of the

natural frequencies is expanded to a larger interval $[0, 30]$ so that synchronization is not reached, our algorithm is able to dig out the coupling strengths as shown in figure 16(C). To quantify the performance of the algorithm at different levels of synchronization, we employ two statistical indicators—the order parameter $0 \leqslant R \leqslant 1$ defined in equation (17) and the G-P dimension [66] for the oscillator network. The order parameter reflects the degree of synchronization, which is between 0 (fully random) and 1 (fully synchronous). The fractal dimension gives the amount of entropy contained in the time series. The larger the value, the more entropy the time series contains.

When the natural frequency is distributed in $[0, 1]$, $< R >= 0.97$ as given in figure 16(D), which is close to a full synchronization. Accordingly, the G-P dimension is only 1.45, indicating that the time series contains very little variability besides the nearly periodic oscillation, which makes it hard to distinguish different nodes. However, the entropy contained in the time series considerably increases if the frequency distribution is extended to the interval $[0, 30]$, since the order parameter and the G-P dimension turn 0.38 and 5.2 respectively as plotted in figure 16(E), signalling little synchronization in the network. To see clearly the influence of heterogeneity induced by the frequency distribution $[0, \mu]$, we scan different values of μ and depict the results in figure 16(F). The order parameter of the system decreases while the normalized dimensionality gradually increases, which indicates the waning of the synchronization. As expected, the accuracy of the reconstruction is gradually increasing and approaching 100%.

3. Discussion

Reconstructing system dynamics and interaction topology is an important yet challenging problem, especially with only partial information of the system state and in the presence of noise. In this article, we proposed a new framework for the reconstruction based on minimization of a carefully designed cost functional. To proceed in the gradient-descent direction in a continuous manner, a new system of differential equations are derived which serve to drive both the unknown observables and parameters to the correct values with the help of given time series. Several examples including the coupled Lorenz system, the FHN neural networks, the coupled Kuramoto oscillators and a gene regulation network (appendix figure 20) are used to demonstrate validity of the current technique. Because of the evolution character, the method could be conveniently utilized to reconstruct complex dynamics and equations with nonlinear parameter dependence, or to monitor system reconfiguration changes on the scene.

With small noise in the system, the method is still useful while a direct application may fail in the presence of strong noise. So, it would be very interesting and helpful to extend the current frame to treat systems contaminated with strong noise. One possible way is to suppress noise in the available data with various existing techniques, for example, by using Wang's method [43] to give a clean orbit first. Nevertheless, strong noise could even move the average orbit, which may be regarded as a renormalizing effect originating from the ignored degrees of freedom. It is desirable to utilize the framework here to renormalize complex dynamics with only slow degrees of freedom that are of our interest, treating those fast degrees as noise.

For small networks, a full connection assumption is made during the reconstruction and then the true network structure and connection strength are conveniently deduced with our method. But for large networks, prior knowledge of network topology is still needed since the number of connections grows quadratically with the network size while the available information increases only linearly. In practice, most networks are sparse and various techniques exist to infer possible connections. Even if the inferred links are spurious due to indirect interaction or confounding effects, our method still works. Nevertheless, it should be of great practical interest to extend the current framework to directly treat large-scale reconstruction without resorting to other inference programs beforehand.

In the current work, the form of equations of motion is assumed known and only unknown parameters need to be deduced. Compared to existing algorithms [30, 44, 67, 68], this is not a serious restriction since we can always extend the expansion basis on the right hand side of the equation to include possible new terms. The cost is that more parameters should be deduced during the reconstruction, which is not a problem if the number of parameters is not over large. Of course, it remains a big challenge to deduce the underlying laws of motion purely from observation without much prior assumption although some progress has been made with machine learning [69–72]. It would be very interesting if we could combine the current framework with these cutting-edge techniques.

If the given information is too little, it seems impossible to deduce much as expected in general. Synchronization effectively reduces the fractal dimensions of the time series and makes it hard to carry out reconstruction, which also brings a lot of interesting problems. To get out of synchronization, we may consider transient dynamics or supply noise to unfold the reduced dimension. If it is impossible to apply

external perturbation, it may be interesting to design an effective model to describe the reduced dynamics as indicated above. If synchronization is partial as in most cases, it is essential to determine and measure typical nodes to enable an equivalent modelling. In this case, some nodes may contain more information than others, how to find these is also an intriguing problem.

In all, from the given information, how much we can deduce about the dynamics and structure of a system is an interesting problem both in theory and in application. We hope that the presented framework in the current paper will provide an alternative route for further investigation.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Acknowledgments

Conflict of interest

The authors declare that they have no conflict of interest.

Author contributions statement

Yueheng Lan, Lili Gui and Kun Xu conceived the research, Yueheng Lan and Zishuo Yan designed the research, Zishuo Yan performed the research, Zishuo Yan and Yueheng Lan analyzed the results, Zishuo Yan and Yueheng Lan wrote this manuscript. All authors reviewed the manuscript.

Additional information

All data, models, generated or used during the study appear in the submitted article, code generated during the study are available from the corresponding author by request.

Appendix

The 8-Lorenz system
Here we choose an Erdős-Rényi (ER) network of 8 nodes as an example, depicted in figure 17(A). The coupling strengths between nodes are randomly selected from $\{0.1, 0.2, 0.3\}$. For small networks with unknown structures, we may assume full connection and identify true links based on the construction result. The inference of coupling strengths with our method is shown in figure 17(B). It can be seen from the figure that with the evolution of inference equations, the coupling strengths change greatly from the initial value 1 at the very start, and converge to the correct values after the transient. When searching in a huge parameter space, the parameters are wandering a lot at the beginning and the error function Δ is not always decreasing (appendix figure 19) due to the state change over time. Nevertheless, once the parameters enter the basin of attraction, it will be quickly attracted to the correct values which minimize Δ. Very small inferred strengths indicate no connection between the corresponding nodes. In figures 17(C) and (D) the parameters of one node and the reconstructed time series of the variable y and z are displayed. Just like a single Lorenz system, the inferred results are quite accurate.

Recycling of data
In our reconstruction process, the amount of data needed is very large. For example, for the Lorentz system we need 50 000 sampling points to recover the dynamics. What if there are not so many sampling points? Our
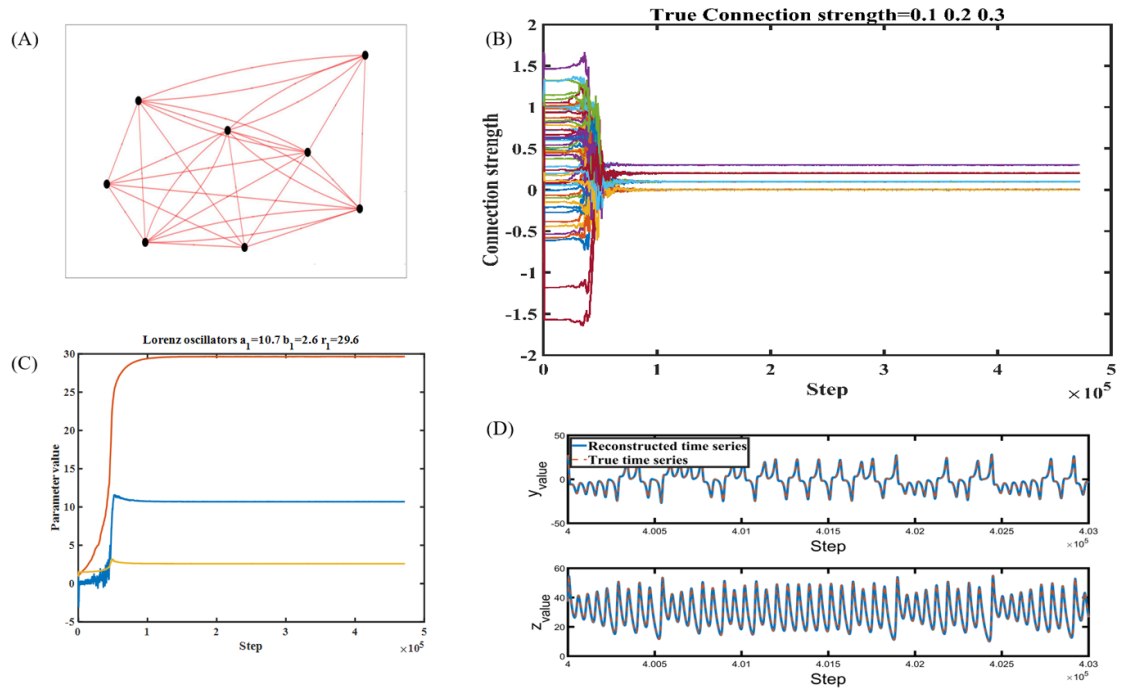
Figure 17. Reconstruction in a network of 8 coupled Lorenz systems with different local parameters. (A) The coupling network. (B) Inference of the coupling strengths. (C) Local parameter reconstruction of one node. (D) Reconstruction of the time series of the hidden variable y and z of that node.
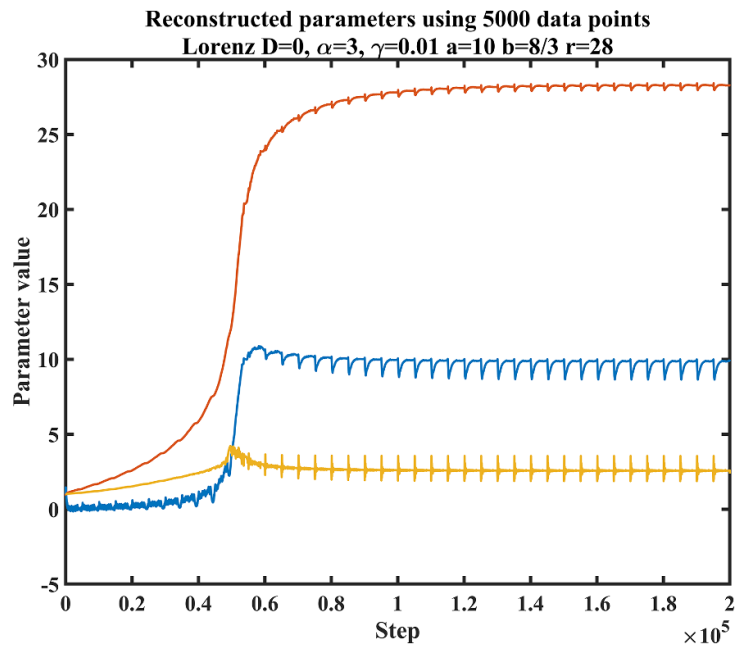
Figure 18. Reconstruct the Lorentz system by recycling 5000 data points.

solution is to recycle the measured data to generate longer orbits. As shown in figure 18, we only used 5000 data points for the reconstruction. It seems that all the unknown parameters converge to the right values but with jittering appearing every 5000 points. This is because the recycled orbit breaks every 5000 points and so the continuous evolution is restarted with a jump at this point. Nevertheless, the quick relaxation of parameter values before next jump is clearly seen, which indicates the robustness of the reconstruction and feasibility of recycling the data.
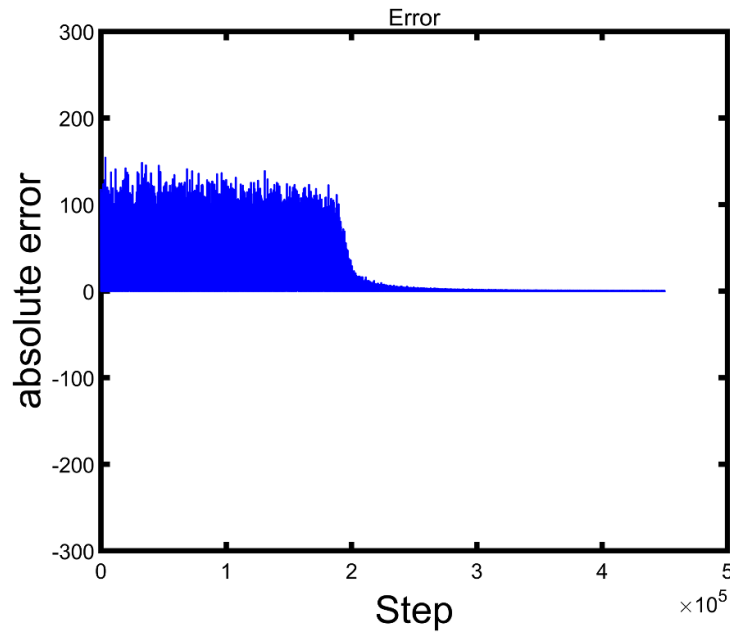
Figure 19. The error function Δ in the 8-Lorenz system.

Error function Δ in equation (3)

Although the inference system evolves along the negative gradient, the error function Δ defined in equation (3) does not always decrease since it is explicitly a time-dependent function. As shown in figure 19, at the initial stage, the error Δ indeed fluctuates a lot. But after entering the basin of attraction of the parameters and the dynamics, the convergence dominates the time evolution so that the error suddenly drops to almost 0 as seen in the figure.

Gene regulation dynamics

Due to its evolution character, the current algorithm may treat dynamics with a nonlinear parameter dependence. To demonstrate this salient point, we use a simple gene regulatory network [73].
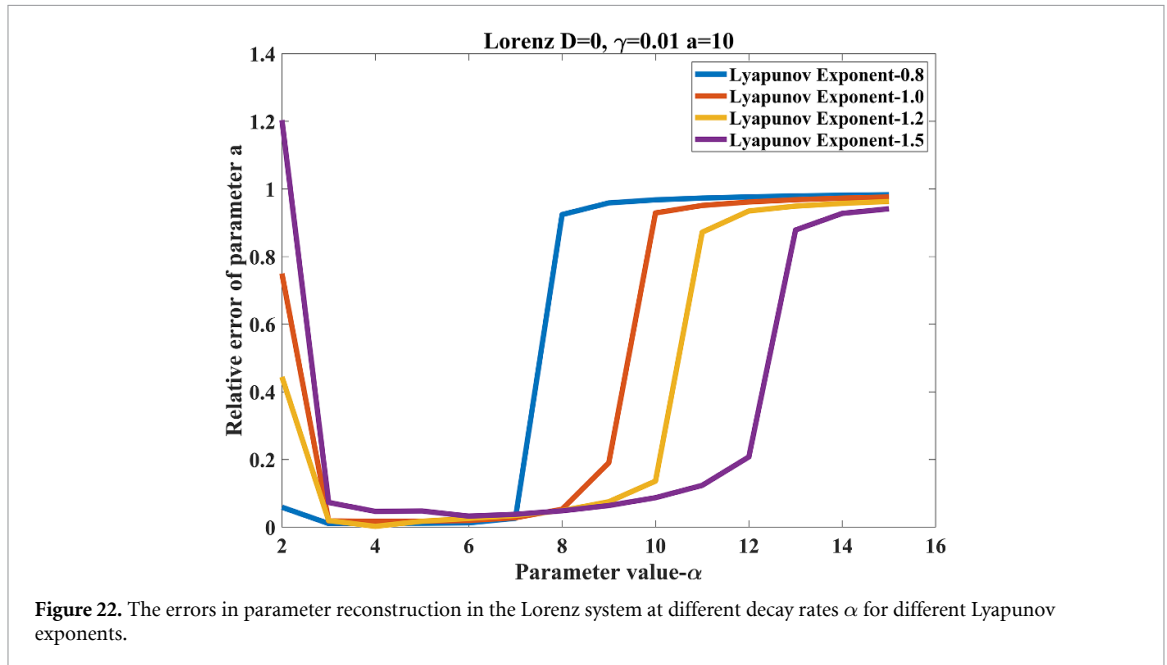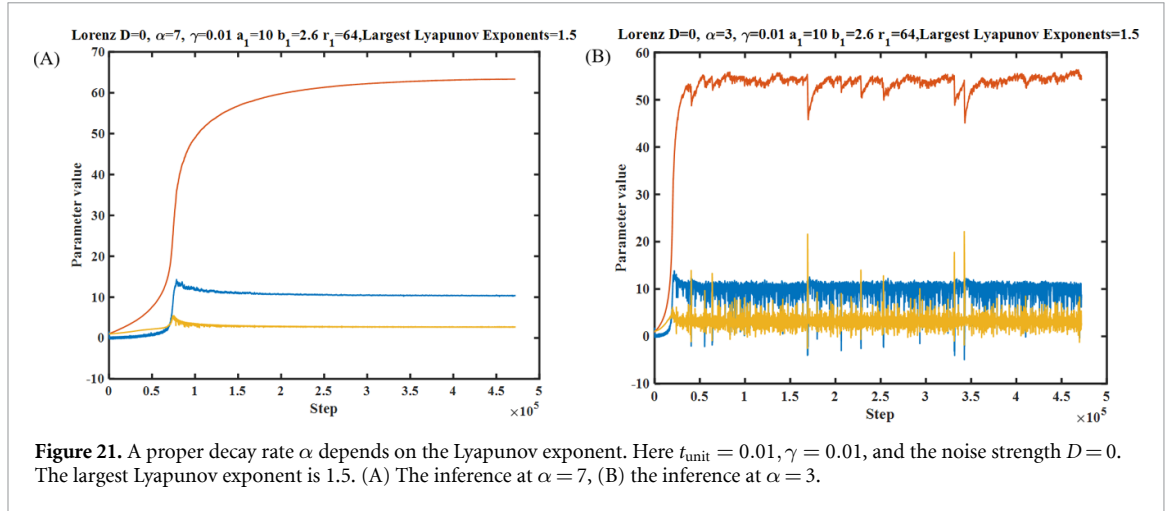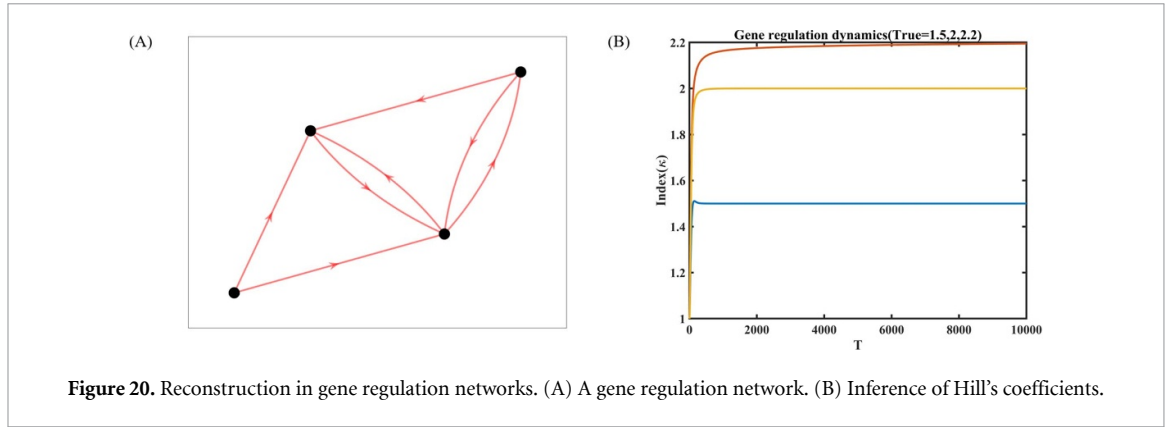
$$\frac{dg_i}{dt} = a_i g_i + \sum_{j=1}^{N} A_{i,j} \frac{g_j^{\kappa_j}}{g_j^{\kappa_j} + 1} \tag{18}$$

where g_i is the concentration of gene i, κ_i denotes the Hill's coefficients, and $a_i \in [-0.3, -0.1]$. Parameters $A_{i,j}$ denotes the regulation strength of gene j on gene i which is set to 0.1 and the network is plotted in figure 20(A). Given the time series of g_i, we would like to deduce the connection A_{ij} and the Hill's coefficients in the equation, which take three different values. To start the reconstruction, all the values of the parameters are initially set to 1 in the inference equation. The inference dynamics is depicted in figure 20(B). It is easy to see that the unknown parameters are quickly driven to the true values, although they appear as power indicies in equation (18). Hence, equipped with the current reconstruction technique, it seems possible to tackle intricate dynamics with nontrivial parameter dependence in real-world applications.

Lyaponov exponent *vs* the decay rate α

The choice of the decay rate α is closely related to the Lyapunov exponent. To compare figures 2 and 3, we increase the parameter r in the Lorentz system (10) from 28 to 64, and as a result the largest Lyapunov exponent increases from 0.8 to 1.5. If we still take the value 3 of the decay rate α as in figure 2, the inference fluctuates a lot with no perceivable convergence as depicted in figure 21(B). In fact, to ensure the convergence, a larger $\alpha = 7$ is needed to counter the increased instability as shown in figure 21(A).

In figure 22, a more comprehensive comparison is made at different decay rates for different Lyapunov exponents. The effective α tends to increase with the Lyapunov exponent, being consistent with the above observation. It is good to note that the increase is slow. There seems exist an interval such that the α in this interval is generally good for different α values.

Figure 20. Reconstruction in gene regulation networks. (A) A gene regulation network. (B) Inference of Hill's coefficients.

Figure 21. A proper decay rate α depends on the Lyapunov exponent. Here $t_{unit} = 0.01, \gamma = 0.01$, and the noise strength $D = 0$. The largest Lyapunov exponent is 1.5. (A) The inference at $\alpha = 7$, (B) the inference at $\alpha = 3$.

Figure 22. The errors in parameter reconstruction in the Lorenz system at different decay rates α for different Lyapunov exponents.

FHN synchronized dynamics

Synchronization generally makes the inference harder, which can be clearly seen here from a simple example. As shown in figure 23(A), the synchronization dynamics of the three FHN nodes are used in the reconstruction. When node 2 and 3 are fully synchronized, it is impossible to distinguish them only from the time series of node 1 (figure 23(B)). Therefore, in this case, the obtained coupling strength 5 is the average of the original two: 4 and 6. If they are not fully synchronized (by changing the parameter values), our method works but is slow to infer the correct coupling strength when close to synchronization (figure 23(C).

Figure 23. Reconstruction against synchronization in a network with three FHN nodes. (A) The network with the coupling strength $= [2.5\ 2.5\ 4\ 6]$, $a_i = 0.3, b_i = 0.5, c_i = -0.04$. Inference of the coupling strength in full synchronization (B) and in partial synchronization (C), with $a_1 = 0.3, a_2 = 0.4, b_1 = 0.5, b_2 = 0.5, c_1 = -0.04, c_2 = -0.03$.
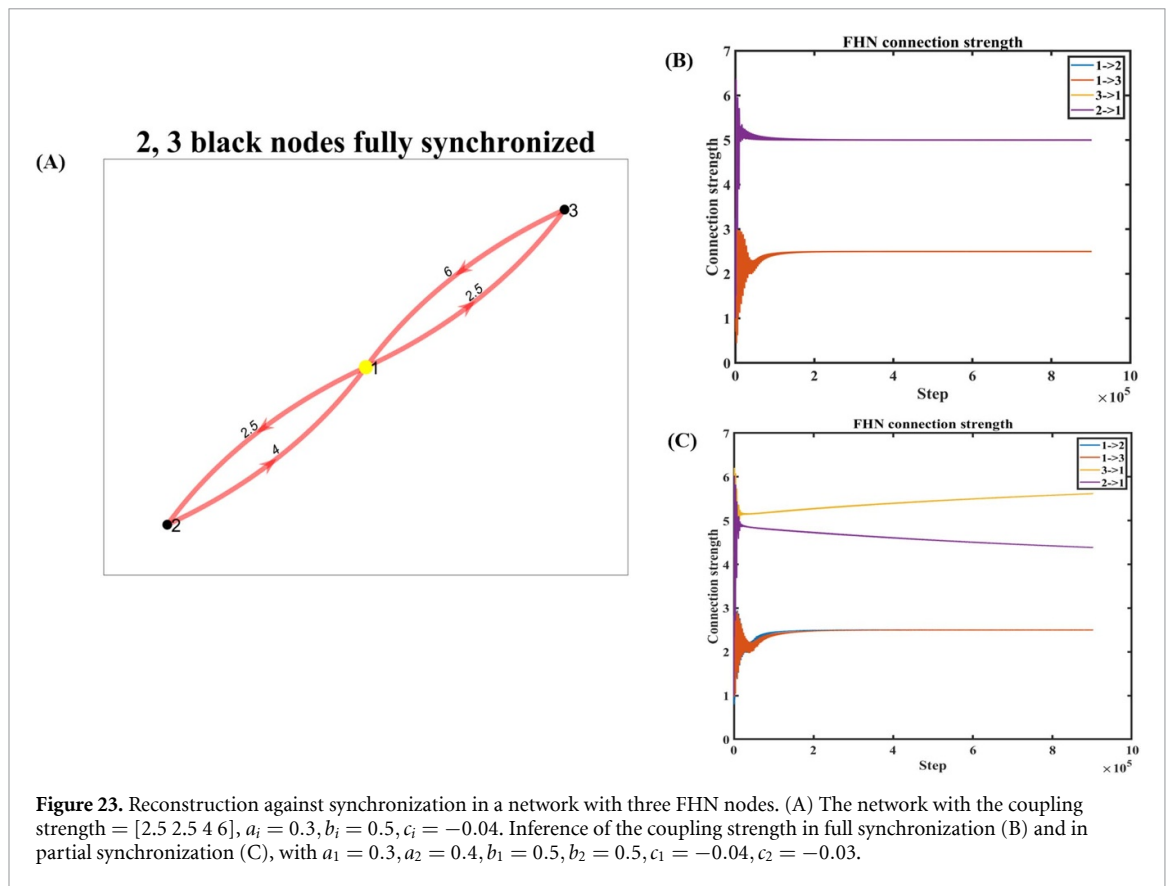
References

[1] Schütte R and Zelewski S 1992 *Nonlinear Modeling and Forecasting* (Addison-Wesley Publishing Company)
[2] Winkel P 1995 Application of time series analysis in the clinical setting *Scand. J. Clin. Lab. Invest.* **55** 11–16
[3] Pastén D, Czechowski Z and Toledo B 2018 Time series analysis in earthquake complex networks *Chaos* **28** 083128
[4] Gouveia N 2000 Time series analysis of air pollution and mortality: effects by cause, age and socioeconomic status *J. Epidemiol. Commun. Health* **54** 750–5
[5] Caldarelli G, Chessa A, Pammolli F, Gabrielli A and Puliga M 2013 Reconstructing a credit network *Nat. Phys.* **9** 125–6
[6] Donges J F, Zou Y, Marwan N and Kurths J 2009 The backbone of the climate network *Europhys. Lett.* **87** 0295–5075
[7] Chen G R 2014 The china power grid: a network science perspective *Natl Sci. Rev.* **1** 368
[8] C. elegans Sequencing Consortium 1998 Genome sequence of the nematode C. elegans: a platform for investigating biology *Science* **282** 2012–8
[9] Gardner T S, Bernardo D D, Lorenz D and Collins J J 2003 Inferring genetic networks and identifying compound mode of action via expression profiling *Science* **301** 102–5
[10] Margolin A, Nemenman I and Basso K 2006 An algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context *BMC Bioinform.* **1** s1–7
[11] Eagle N, Pentland A and Lazer D 2009 Inferring friendship network structure by using mobile phone data *Proc. Natl Acad. Sci. USA* **106** 15274–8
[12] Stetter O, Battaglia D, Soriano J, Geisel T and Beggs J 2012 Model-free reconstruction of excitatory neuronal connectivity from calcium imaging signals *Plos Comput. Biol.* **8** 1002653
[13] Palsson B 2015 *Systems Biology- Properties of Reconstructed Networks* (Cambridge University Press)
[14] Boccaletti S, Latora V, Moreno Y, Chavez M and Hwang D U 2006 Complex networks: structure and dynamics *Phys. Rep.* **424** 175–308
[15] Hufnagel L, Brockmann D and Geisel T 2004 Forecast and control of epidemics in a globalized world *Proc. Natl Acad. Sci. USA* **101** 15124–9
[16] Bullmore E and Sporns O 2009 Complex brain networks: graph theoretical analysis of structural and functional systems *Nat. Rev. Neurosci.* **10** 186–98
[17] Breakspear M 2017 Dynamic models of large-scale brain activity *Nat. Neurosci.* **20** 340–52
[18] Newman M 2003 The structure and function of complex networks *SIAM Rev.* **45** 167–256
[19] Timme M 2007 Revealing network connectivity from response dynamics *Phys. Rev. Lett.* **98** 224101
[20] Shandilya S G and Timme M 2011 Inferring network topology from complex dynamics *New J. Phys.* **13** 87–92
[21] Napoletani D and Sauer T D 2008 Reconstructing the topology of sparsely connected dynamical networks *Phys. Rev.* E **77** 026103
[22] Brunton SL, Proctor J L and Kutz J N 2016 Discovering governing equations from data by sparse identification of nonlinear dynamical systems *Proc. Natl Acad. Sci. USA* **113** 3932–7
[23] Casadiego J, Nitzan M, Hallerberg S and Timme M 2017 Model-free inference of direct network interactions from nonlinear collective dynamics *Nat. Commun.* **8** 1–10
[24] Lorenz E N 1991 Dimension of weather and climate attractors *Nature* **353** 241–4
[25] Frenzel S and Pompe B 2007 Partial mutual information for coupling analysis of multivariate time series *Phys. Rev. Lett.* **99** 204101

[26] Pompe B and Runge J 2011 Momentary information transfer as a coupling measure of time series *Phys. Rev.* E **83** 051122

[27] Schreiber T 2000 Measuring information transfer *Phys. Rev. Lett.* **85** 461–4

[28] Marinazzo D, Pellicoro M and Stramaglia S 2008 Kernel method for nonlinear granger causality *Phys. Rev. Lett.* **100** 144103

[29] Baake E, Baake M, Bock H G and Briggs K M 1992 Fitting ordinary differential equations to chaotic data *Phys. Rev.* A **45** 5524–9

[30] Rudy S H, Brunton S L, Proctor J L and Kutz J N 2017 Data-driven discovery of partial differential equations *Sci. Adv.* **3** e1602614

[31] Yu D, Righero M and Kocarev L 2006 Estimating topology of networks *Phys. Rev. Lett.* **97** 188701

[32] Zhou J and Lu J A 2007 Topology identification of weighted complex dynamical networks *Physica* A **386** 481–91

[33] Wu X 2008 Synchronization-based topology identification of weighted general complex dynamical networks with time-varying coupling delay *Physica* A **387** 997–1008

[34] Liu H, Lu J A, Lü J and Hill D J 2009 Structure identification of uncertain general complex dynamical networks with time delay *Automatica* **45** 1799–807

[35] Parlitz U 1996 Estimating model parameters from time series by autosynchronization *Phys. Rev. Lett.* **76** 1232–5

[36] Tao C, Zhang Y and Jiang J J 2007 Estimating system parameters from chaotic time series with synchronization optimized by a genetic algorithm *Phys. Rev.* E **76** 016209

[37] Levnaji Z and Pikovsky A 2011 Network reconstruction from random phase resetting *Phys. Rev. Lett.* **107** 34101

[38] Ren J, Wang W X, Li B and Lai Y C 2010 Noise bridges dynamical correlation and topology in coupled oscillator networks *Phys. Rev. Lett.* **104** 058701

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

[40] Ching E and Tam H C 2017 Reconstructing links in directed networks from noisy dynamics *Phys. Rev.* E **95** 010301

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

[42] Yang C, Zhang Z, Chen T, Wang S and Hu G 2017 Reconstruction of noise-driven nonlinear networks from node outputs by using high-order correlations *Sci. Rep.* **7** 44639

[43] Wang J, Yan Z, Gui L, Xu K and Lan Y 2022 Reconstruction of nonlinear flows from noisy time series *Nonliner Dyn.* **108** 3887–902

[44] Gao T T and Yan G 2022 Autonomous inference of complex network dynamics from incomplete and noisy data *Nat. Comput. Sci.* **2** 160–8

[45] Breeden J L and Hübler A 1990 Reconstructing equations of motion from experimental data with unobserved variables *Phys. Rev.* A **42** 5817

[46] Cremers J and Hübler A 1987 Construction of differential equations from experimental data *Z. Naturforsch.* A **42** 797–802

[47] James P, Crutchfield B and McNamara S 1987 Equations of motion from a data series *Complex Syst.* **1** 417–52

[48] Gouesbet G 1991 Reconstruction of standard and inverse vector fields equivalent to a rossler system *Phys. Rev.* A **44** 6264–80

[49] Gouesbet G 1992 Reconstruction of vector fields: the case of the lorenz system *Phys. Rev.* A **46** 1784–96

[50] Wang W X, Lai Y C, Grebogi C and Ye J 2011 Network reconstruction based on evolutionary-game data via compressive sensing *Phys. Rev.* X **1** 021021

[51] Su R Q, Wang W X and Lai Y C 2012 Detecting hidden nodes in complex networks from time series *Phys. Rev.* E **85** 1149–64

[52] Su R Q, Lai Y C, Wang W X and Do Y 2014 Uncovering hidden nodes in complex networks in the presence of noise *Sci. Rep.* **4** 3944

[53] Su R Q, Wang W X, Wang X and Lai Y C 2016 Data-based reconstruction of complex geospatial networks, nodal positioning and detection of hidden nodes *R. Soc. Open Sci.* **3** 150577

[54] Shen Z, Wang W X, Fan Y, Di Z and Lai Y C 2014 Reconstructing propagation networks with natural diversity and identifying hidden sources *Nat. Commun.* **5** 4323

[55] Wu X, Wang W and Zheng W X 2012 Inferring topologies of complex networks with hidden variables *Phys. Rev.* E **86** 046106

[56] Guo S, Seth A K, Kendrick K M, Zhou C and Feng J 2008 Partial granger causality-eliminating exogenous inputs and latent variables *J. Neurosci. Meth.* **172** 79–93

[57] Yang C, Zhang Z, Chen T Y, Wang S H and Gang H 2017 Reconstruction of noise-driven nonlinear dynamic networks with some hidden nodes *Sci. China Phys. Mech.* **60** 46

[58] Zhang Z, Yang C and Gang H 2017 Network reconstructions with partially available data *Front. Phys. China* **12** 117–23

[59] Shi R, Jiang W and Wang S 2020 Detecting network structures from measurable data produced by dynamics with hidden variables *Chaos* **30** 138

[60] Ching E and Tam P H 2018 Effects of hidden nodes on the reconstruction of bidirectional networks *Phys. Rev.* E **98** 062318

[61] Lorenz E N 1963 Deterministic nonperiodic flow *J. Atmos. Sci.* **20** 130–41

[62] Ott E 2002 *Chaos in Dynamical Systems* (Cambridge University Press)

[63] Kuramoto Y 1975 Self-entrainment of a population of coupled non-linear oscillators *Int. Symp. on Mathematical Problems in Theoretical Physics* (Springer) pp 420–2

[64] Pietras B and Daffertshofer A 2019 Network dynamics of coupled oscillators and phase reduction techniques *Phys. Rep.* **819** 1–105

[65] Acebrón J A, Bonilla L L, Vicente C J P, Ritort F and Spigler R 2005 The kuramoto model: a simple paradigm for synchronization phenomena *Rev. Mod. Phys.* **77** 137–85

[66] Grassberger P and Procaccia I 1983 Characterization of strange attractors *Phys. Rev. Lett.* **50** 346

[67] Kera H and Hasegawa Y 2016 Noise-tolerant algebraic method for reconstruction of nonlinear dynamical systems *Nonlinear Dyn.* **85** 675–92

[68] Aguirre L, Freitas U, Letellier C and Maquet J 2001 Structure-selection techniques applied to continuous-time nonlinear models *Physica* D **158** 1–18

[69] Cuomo S, Cola V S D, Giampaolo F, Rozza G, Raissi M and Piccialli F 2022 Scientific machine learning through physics-informed neural networks: where we are and what's next *J. Sci. Comput.* **92** 88

[70] Udrescu S M and Tegmark M 2019 Ai feynman: a physics-inspired method for symbolic regression *Sci. Adv.* **6** eaay2631

[71] Zhang Y, Guo Y, Zhang Z, Chen M, Wang S and Zhang J 2022 Universal framework for reconstructing complex networks and node dynamics from discrete or continuous dynamics data *Phys. Rev.* E **106** 034315

[72] Raissi M, Perdikaris P and Karniadakis G 2019 Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations *J. Comput. Phys.* **378** 686–707

[73] Mazur J, Ritter D, Reinelt G and Kaderali L 2009 Reconstructing nonlinear dynamic models of gene regulation using stochastic sampling *BMC Bioinform.* **10** 448