



PERSPECTIVE

# Reconstructing networks via discrete state dynamical data: A mini-review

To cite this article: Chuang Ma *et al* 2023 *EPL* **144** 21002

View the [article online](#) for updates and enhancements.

## You may also like

- [Interlacing relaxation and first-passage phenomena in reversible discrete and continuous space Markovian dynamics](#)  
David Hartich and Aljaž Godec
- [Denoising non-steady state dynamic PET data using a feed-forward neural network](#)  
G I Angelis, O K Fuller, J E Gillam et al.
- [Hyperaccurate bounds in discrete-state Markovian systems](#)  
D M Busiello and C E Fiore

## Perspective

# Reconstructing networks via discrete state dynamical data: A mini-review

CHUANG MA<sup>1</sup>, HUAN WANG<sup>2</sup> and HAI-FENG ZHANG<sup>3(a)</sup><sup>1</sup> School of Internet, Anhui University - Hefei, 230601, China<sup>2</sup> School of Big Data and Statistics, Anhui University - Hefei, 230601, China<sup>3</sup> School of Mathematical Science, Anhui University - Hefei, 230601, Chinareceived 17 August 2023; accepted in final form 27 October 2023  
published online 13 November 2023

**Abstract** – The inference of network structure from dynamic data is one of the most challenging scientific problems in network science. To address this issue, researchers have proposed various approaches regarding different types of dynamical data. Since many real evolution processes or social phenomena can be described by discrete state dynamical systems, such as the spreading of epidemic, the evolution of opinions, and the cooperation behaviors, network reconstruction methods driven by discrete state dynamical data were also widely studied. In this letter, we provide a mini-review of recent progresses for reconstructing networks based on discrete state dynamical data. These studies encompass network reconstruction problems where the dynamical processes are known, as well as those where the dynamics are unknown, and extend to the reconstruction of higher-order networks. Finally, we discuss the remaining challenges in this field.

perspective

Copyright © 2023 EPLA

**Introduction.** – Since many real-world systems can be abstracted as complex networks with interactions between nodes, such as power grids, transportation networks, social networks, biological networks, and so on. Complex network science has received significant attention in the past few decades. Wherein the analysis of network structures and the investigation of collective dynamics on these structures are the main focuses in network sciences [1–3]. The aforementioned studies are all built upon the assumption that the network structures are known. However, in many real situations, the topology of a network is often difficult to directly obtain, giving rise to the problem of network reconstruction [4,5].

The network reconstruction problem is one of the significant “inverse problems” in network sciences, which involves inferring the connectivity among individuals based on the observation of their dynamical data [4–18], implying great practical values in many fields. One of its most classical applications is the reverse engineering of gene regulatory networks in biology. Gene expression data is assumed to follow a nonlinear dynamical model, and gene regulatory networks are inferred by linearizing the model with small perturbations at steady states [19–24]. Up to now, various network reconstruction approaches

have been developed from different perspectives. For the continuous state dynamical systems, whose dynamical processes are mainly expressed in the form of differential equations, such as reaction-diffusion, consensus, and synchronization. Many methods were developed to reconstruct networks from such dynamical data, including perturbation approaches [19–21], noise-driven approaches [25–28], driving response approaches [29–31], adaptive control approaches [32–35], error function minimization approaches [11,18,36–38], machine learning approaches [39,40], and so forth. In addition, some methods can be extended to reconstruct other types of networks, *e.g.*, weighted networks [41,42] and multi-layer networks [43,44], to reconstruct dynamical systems and predict system catastrophes [13], or to solve network transport problems [45].

Like the continuous state dynamical systems, discrete state dynamical systems are also widespread in nature and human society, examples include infection and susceptibility in epidemic spreading, cooperation and defection in evolutionary game, approval and opposition in voting events, and the spin-up and spin-down states of atoms in ferromagnetic materials. These discrete dynamical models can be generalized to complex networks, such as the Susceptible-Infectious-Susceptible (SIS) model on networks [46], game model on networks [47,48], voter

<sup>(a)</sup>E-mail: haifengzhang1978@gmail.com (corresponding author)

$$P\{\{s_{t_m+1}^i\}_{m=1,\dots,M}|\{s_{t_m}^j\}_{j=1,\dots,N;m=1,\dots,M},\{a_{i,j}\}_{j=1,\dots,N},\Omega\} = \prod_m [(P_{0 \rightarrow 1}^{i,t_m})^{(1-s_{t_m}^i)s_{t_m+1}^i} (1 - P_{0 \rightarrow 1}^{i,t_m})^{(1-s_{t_m}^i)(1-s_{t_m+1}^i)} (P_{1 \rightarrow 0}^{i,t_m})^{s_{t_m}^i(1-s_{t_m+1}^i)} (1 - P_{1 \rightarrow 0}^{i,t_m})^{s_{t_m}^i s_{t_m+1}^i}], \quad (5)$$

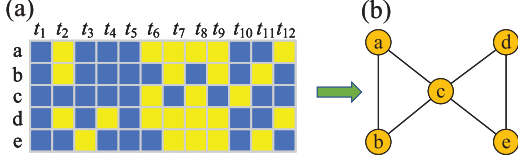


Fig. 1: Schematic of network reconstruction based on the discrete state dynamical data. (a) Binary-state time series, with different colors representing two distinct states. (b) The target network structure to be reconstructed.

model on networks [49], Ising model on networks [50], and so forth. In many situations, one can only observe the data generated by the discrete state dynamical models hosting on complex networks, but the structures of the networks are unknown. Thus, we should develop some methods to reconstruct the networks based on these discrete state dynamical data.

Specifically, the network reconstruction problem based on discrete state dynamical data, as shown in fig. 1, can be addressed by observing the state sequences of nodes (*e.g.*, whether they are infected with a disease or have received information) to reconstruct the structures of networks where disease or information propagation takes place. Assuming  $s_t^i$  represents the state of node  $i$  at time  $t$ , here we primarily consider binary-state data, where  $s_t^i = 1$  indicates that node  $i$  is in an active state at time  $t$ , such as the infected state in disease spreading models or the spin-up state in the Ising model, etc. Conversely,  $s_t^i = 0$  indicates that node  $i$  is in an inactive state at time  $t$ , such as the susceptible state in disease spreading models or the spin-down state in the Ising model, etc. The network reconstruction problem from discrete state dynamical data involves inferring the network's topological structure from the observed node state data  $s_t^i$  ( $i = 1, 2, \dots, N; t = 1, 2, \dots, M$ ), where  $N$  represents the number of nodes, and  $M$  represents the length of the time series.

**Network reconstruction with known node dynamics.** – When the mechanism behind the time series data is known, *i.e.*, the dynamical model on networks is known, the transition probabilities  $P_{0 \rightarrow 1}^{i,t}$  and  $P_{1 \rightarrow 0}^{i,t}$  for node  $i$  at time  $t$  from state 0 to state 1 and from state 1 to state 0 can be calculated. As for the SIS dynamics, the probability that node  $i$  transitions from the susceptible state (*i.e.*, 0) to the infected state (*i.e.*, 1) at time  $t_m$  is

$$P_{0 \rightarrow 1}^{i,t_m} = 1 - (1 - \lambda^i)^{\sum_{j \neq i} s_{t_m}^j a_{i,j}}. \quad (1)$$

The probability that node  $i$  transitions from the infected state to the susceptible state is

$$P_{1 \rightarrow 0}^{i,t_m} = \mu^i, \quad (2)$$

where  $\lambda^i$  and  $\mu^i$  are the infection rate and recovery rate of node  $i$ , respectively.  $a_{i,j} = 1$  indicates that node  $j$  is a neighbor of node  $i$ , and  $a_{i,j} = 0$  indicates that node  $j$  is not a neighbor of node  $i$ . Equation (2) indicates that, for SIS dynamics, the transition probability from the infected state to the susceptible state does not depend on the neighbors' states, namely, it is not useful to infer the network structure. As a result, we only need to focus on eq. (1), and it can be transformed as follows [14]:

$$\ln[1 - P_{0 \rightarrow 1}^{i,t_m}] = \ln(1 - \lambda^i) \sum_{j \neq i} s_{t_m}^j a_{i,j}. \quad (3)$$

The values of  $P_{0 \rightarrow 1}^{i,t_m}$  and  $s_{t_m}^j$  can be calculated from the time series data, and the network reconstruction problem is transformed into solving the following system of linear equations:

$$C^i \mathbf{A}^i = \mathbf{Y}^i, \quad (4)$$

where  $\mathbf{A}^i = [\ln(1 - \lambda^i) a_{i,1}, \dots, \ln(1 - \lambda^i) a_{i,i-1}, \ln(1 - \lambda^i) a_{i,i+1}, \dots, \ln(1 - \lambda^i) a_{i,N}]^T$ ,  $C^i$ , and  $\mathbf{Y}^i$  are known. Due to the sparsity of  $\mathbf{A}^i$ , the compressive sensing method [51,52] is used to solve eq. (4), which allows us to infer the neighbors of node  $i$ . Once the neighbors of all nodes are inferred, it is possible to reconstruct the entire network's structure.

Unlike the SIS dynamics described in eq. (4), not all nodal dynamics equations can be written in the form of linear equations concerning network structure. In this case, the compressive sensing based method cannot be used to reconstruct networks for these discrete state dynamics. However, for the most discrete state dynamics, the transition probabilities depend on the states of neighboring nodes, *i.e.*,  $P_{0 \rightarrow 1}^{i,t}$  and  $P_{1 \rightarrow 0}^{i,t}$  are the functions of  $\sum_{j \neq i} s_t^j a_{i,j}$ . Based on this fact, to reconstruct the neighbors of node  $i$ , the following likelihood function can be written [53,54]:

*see eq. (5) above*

where  $\Omega$  represents dynamical parameters such as the infection rate and recovery rate in SIS dynamics.

Let us take the logarithmic form of eq. (5) and differentiating it with respect to  $\{a_{i,j}\}_{j=1,\dots,N}$ , leading to a high-dimensional nonlinear system of equations. Then

a mean-field approximation is applied to eq. (5), namely,

$$\frac{\sum_{j \neq i} s_t^j a_{i,j}}{k^i} \approx \frac{\sum_{j \neq i} s_t^j}{N-1}, \quad (6)$$

where  $k^i$  represents the degree of node  $i$ . As a result, the high-dimensional nonlinear system of equations can be linearized around point  $\frac{k^i}{N-1} \sum_{j \neq i} s_t^j$  with respect to  $\sum_{j \neq i} s_t^j a_{i,j}$  by using the first-order Taylor expansion, and a similar system of linear equations as eq. (4) can be constructed. This method has also been applied to the reconstruction of signed networks [55] and duplex networks [53].

In addition, for infection cascades or information cascades data, which can be simulated by using SI, SIR, or independent cascade (IC) models. It is worth noting that the recovered state (R state) of the SIR model at time  $t$  is denoted as  $s_t^i = 2$ . Therefore, when we have given the network (*i.e.*,  $A$ ) and model parameters (*i.e.*,  $\Omega$ ), the likelihood function is given as follows:

$$P\left\{\{s_{t_m}^i\}_{i=1,2,\dots,N;m=1,2,\dots,M} | A, \Omega\right\}. \quad (7)$$

From Bayes' rule, one can obtain the distribution of interest:

$$\begin{aligned} P\left\{A, \Omega \mid \{s_{t_m}^i\}_{i=1,2,\dots,N;m=1,2,\dots,M}\right\} = \\ \frac{P\left\{\{s_{t_m}^i\}_{i=1,2,\dots,N;m=1,2,\dots,M} | A, \Omega\right\} P\{A, \Omega\}}{P\left\{\{s_{t_m}^i\}_{i=1,2,\dots,N;m=1,2,\dots,M}\right\}}. \end{aligned} \quad (8)$$

To solve eq. (8), the Metropolis-Hastings MCMC algorithm [56] or belief propagation approach [57] can be employed, which allows us to infer the network structure.

**Network reconstruction with unknown node dynamics.** – In fact, it is difficult to precisely know the dynamical equations, that is to say, the transition probability is hard to precisely determine. To solve the problem, Li *et al.* assumed that the transition of a node from an inactive state to an active state is the linear function of the number of its active neighbors [58], that is,

$$P_{0 \rightarrow 1}^{i,t} = c^i \sum_{j \neq i} s_t^j a_{i,j} + d^i. \quad (9)$$

Then a system of linear equations is constructed by utilizing time series data and eq. (9), and the network can be reconstructed by using the compressive sensing method [58]. Since eq. (9) neglects the nonlinear characteristics of transition probability regarding many dynamics, leading to the reconstruction performance is unsatisfactory in some cases. On the one hand, the sigmoid function can well reflect the nonlinear characteristics of transition probability, on the other hand, the sigmoid function can approximately fit many types of functions by varying its parameters. Inspired by the above reasons, we used the sigmoid functions

in logistic regression to approximate the transition probabilities  $P_{0 \rightarrow 1}^{i,t}$  and  $P_{1 \rightarrow 0}^{i,t}$ , and then the likelihood function concerning eq. (5) can be obtained. Finally, the network structure can be inferred by similarly utilizing the above process. In particular, our method can infer the dynamical mechanism of different dynamics by learning the parameters in the sigmoid function [54].

Another reasonable assumption is that, for the discrete binary-state dynamics, the transition probability of a node from an inactive state at time  $t$  to an active state at time  $t+1$  is influenced by its active neighbors at time  $t$ . Therefore, according to the conditional probability formula, one has [59]

$$\begin{aligned} P(s_{t+1}^i = 1, j \rightarrow i | s_t^j = 1, s_t^i = 0) = \\ P(j \rightarrow i | s_t^j = 1, s_t^i = 0, s_{t+1}^i = 1) \\ * P(s_{t+1}^i = 1 | s_t^j = 1, s_t^i = 0), \end{aligned} \quad (10)$$

where  $j \rightarrow i$  represents the direct influence of node  $j$  on node  $i$ . Let  $P_{j \rightarrow i} = P(j \rightarrow i | s_t^j = 1, s_t^i = 0, s_{t+1}^i = 1)$ , and  $P_{j \rightarrow i} > 0$  indicates that node  $j$  is a neighbor of node  $i$ . Further, let  $P_j^i = P(s_{t+1}^i = 1 | s_t^j = 1, s_t^i = 0)$ , and its value can be estimated through the time series data and is a known value. Then, if node  $i$  is not activated at time  $t_m$ , the expected number of times node  $i$  is activated at time  $t_m + 1$  can be expressed as

$$E_i^{t_m+1} = \sum_{j(j \neq i)} P_{j \rightarrow i} P_j^i \Psi_j^{t_m} + \varepsilon_i, \quad (11)$$

where  $\Psi_j^{t_m}$  represents the number of times node  $j$  is activated at time  $t_m$ . For the discrete binary-state dynamics, the node's state can only be in active (*i.e.*, 1) or inactive (*i.e.*, 0) state, *i.e.*,  $\Psi_j^{t_m}$  takes values 1 or 0.  $\varepsilon_i$  represents the reconstruction noise. Then, by assuming that the times that node  $i$  is activated follows a Poisson distribution, and the likelihood function can be formulated by using the time series data. Finally, the EM (Expectation-Maximization) algorithm is applied to solve the value of  $P_{j \rightarrow i}$ , and the entire network structure can be inferred, the whole process is illustrated in fig. 2. This method has also been applied to reconstruct weighted networks [60].

In reality, there are many binary-state dynamics where the two states are not asymmetric (*e.g.*, the SIS dynamics): inactive and active. For instance, in the Ising model, the two states do not have a primary and secondary distinction, *i.e.*, the reversal of a node's state depends on the neighbors with the opposite state. Therefore, any state change of a node provides useful information for network reconstruction. Forcing the node's states to be strictly active and inactive would result in wasting a half of the data, leading to a decrease in the reconstruction accuracy. Assume that a node  $i$  is in state  $\bar{c}_t$  at time  $t$  (one of the two states), and it may be in state  $c_t$  (the other state) at time  $t+1$  owing to the influence of the neighbors with

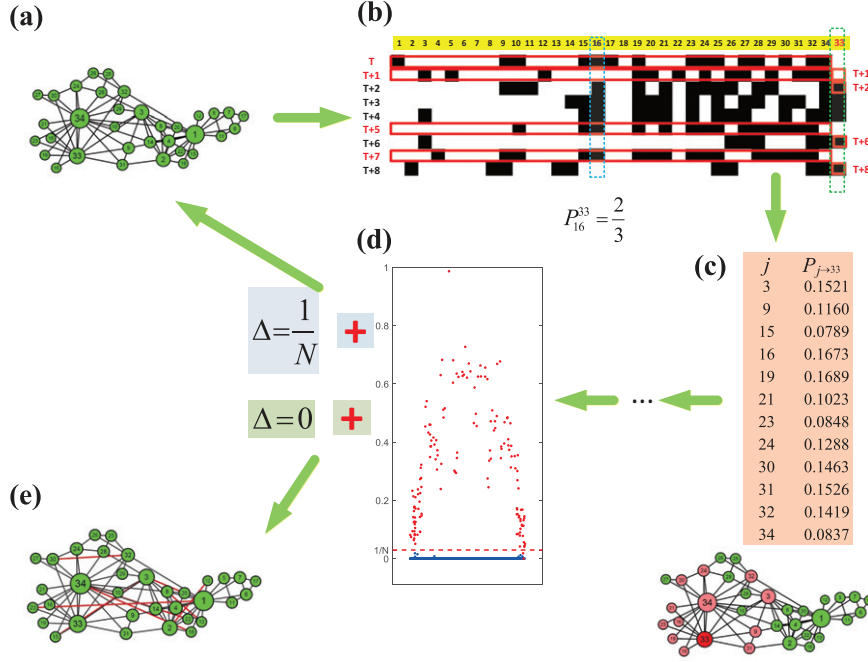


Fig. 2: Network reconstruction illustration. (a) Original network. (b) Discrete time series data. (c) Reconstruction of node 33. (d) The neighbors of each node are inferred in a similar way. (e) Network structure is reconstructed by setting a threshold. This figure is reproduced from [59].

state  $c_t$ . Thus, according to the conditional probability formula, we have [61]

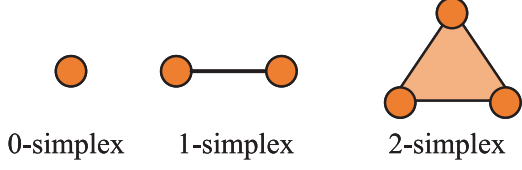
$$\begin{aligned}
 &P(s_{t+1}^i = c_t, j \rightarrow i | s_t^j = c_t, s_t^i = \bar{c}_t) = \\
 &P(j \rightarrow i | s_t^j = c_t, s_t^i = \bar{c}_t, s_{t+1}^i = c_t) \\
 &\quad * P(s_{t+1}^i = c_t | s_t^j = c_t, s_t^i = \bar{c}_t).
 \end{aligned} \tag{12}$$

Here  $P_{j \rightarrow i} = P(j \rightarrow i | s_t^j = c_t, s_t^i = \bar{c}_t, s_{t+1}^i = c_t)$ , and  $P_{j \rightarrow i} > 0$  represents node  $j$  being a neighbor of node  $i$ . Further, let  $P_j^i = P(s_{t+1}^i = c_t | s_t^j = c_t, s_t^i = \bar{c}_t)$ , and its value can be estimated from time series data, too. Then, the expected number of times node  $i$  is in one state at time  $t_m$  and in the other state at time  $t_m + 1$  can be expressed as eq. (11). Then the same approach can be used to infer which nodes are the neighbors of node  $i$ . We need to address that we can determine the nature of the dynamical model from the changes in the time series. For example, in the Ising model, where changes in both states are influenced by the neighbors' states, whereas in the SIS model, only change from one state to another is influenced by neighbors, but not vice versa.

Since reconstructing networks from observed time series data can be considered as a “black-box” problem, deep learning or graph neural network techniques have been used to address this issue [62,63]. Typically, the neural network framework for network reconstruction consists of two parts. The first part is the network generator, which generates network structures, and the second part is the node state learner, which utilizes the generated network structure and the previous node states to

predict the future states of nodes. During the training process, the network generator and the state learner are cross-trained. This approach not only accurately reconstructs networks but also enables dynamical prediction, making it suitable for both discrete and continuous state dynamical processes.

**Higher-order network reconstruction.** – As the data structure of real systems becomes increasingly complex, entity interactions of a large variety of complex systems presents many-body features, thus, pairwise interactions between nodes are no longer sufficient to meet the demand [64,65]. For example, in social contagions, higher-order interaction groups with different sizes can strengthen social contagion effects [66,67]. Typically, simplicial complexes can well describe the many-body interactions between nodes [68–70]. A simplicial complex can be viewed as a network composed of different simplices. A  $k$ -simplex  $\sigma$  is a set  $\sigma = [v_0, \dots, v_k]$  composed of  $k + 1$  nodes, describing a  $(k + 1)$ -body interaction. As shown in fig. 3, a 0-simplex is a single node, a 1-simplex is two nodes connected by an edge, and a 2-simplex is a triangle formed by three nodes, and so on. It is important to note that in a simplicial complex  $\mathcal{K}$ , if a simplex  $\sigma \in \mathcal{K}$ , then any subset of  $\sigma$  must also be included in  $\mathcal{K}$ . For example, a 2-simplicial complex  $\mathcal{K}$  is the collection of 0-simplices, 1-simplices, and 2-simplices, and any two nodes from a 2-simplex will form three 1-simplices. A social contagion model that takes into account the reinforcing effect of 2-simplicial complexes has been proposed in ref. [66], where a susceptible node  $i$  can get infection from an infected


 Fig. 3: Illustration of  $k$ -simplex.

neighbor  $j$  through their pairwise interaction  $(i, j)$  with a probability  $\lambda$ . Node  $i$  can also be infected through a 2-simplex  $(i, j, k)$ , where both  $j$  and  $k$  have already been infected, with the probability  $\lambda_{\Delta}$ , and this probability describes a synergistic reinforcement effect.

Therefore, for the social contagion dynamics on a simplicial complex, the transition of a node from an inactive state to an active state is caused by its neighbors in the active state, and these active neighbors can influence the node's state through two-body interaction (1-simplex), three-body interaction (2-simplex), or higher-order interactions. In addition, in simplicial complexes, one basic fact is that higher-order interactions should include lower-order interactions, for example, if two nodes  $j$  and  $k$  form a 2-simplex with node  $i$ , node  $j$  and node  $k$  must be the neighbors of node  $i$ , namely,  $(i, j)$  and  $(i, k)$  are the 1-simplices. Given the above fact, the strategy for reconstructing simplicial complex networks can be divided into multiple steps [9]. Firstly, the “approximate” neighborhood of each node is inferred by using eqs. (10) and (11), noting that the predicted neighbors in this step are not accurate because the three-body interactions have been ignored. Then, the real 1-simplices and 2-simplices can be further inferred based on the “approximate” neighborhood information in the first step. Specifically, the transition probability of a node from an inactive state to an active state, induced by 2-simplices in an active state, can be expressed as

$$\begin{aligned} & P(s_{t+1}^i = 1, jk \rightarrow i | s_t^i = 0, s_t^j = 1, s_t^k = 1) = \\ & P(jk \rightarrow i | s_t^i = 0, s_t^j = 1, s_t^k = 1, s_{t+1}^i = 1) \\ & * P(s_{t+1}^i = 1 | s_t^i = 0, s_t^j = 1, s_t^k = 1) = \\ & P_{jk \rightarrow i} P_{jk}^i, \end{aligned} \quad (13)$$

where  $jk \rightarrow i$  represents the joint influence of nodes  $j$  and  $k$  on node  $i$ , and if there is an influence, it indicates that nodes  $i$ ,  $j$ , and  $k$  form a 2-simplex  $(i, j, k)$ . Let  $P_{jk \rightarrow i} = P(jk \rightarrow i | s_t^i = 0, s_t^j = 1, s_t^k = 1, s_{t+1}^i = 1)$ , and  $P_{jk \rightarrow i} > 0$  indicates that nodes  $i$ ,  $j$ , and  $k$  form a 2-simplex. Then, let  $P_{jk}^i = P(s_{t+1}^i = 1 | s_t^i = 0, s_t^j = 1, s_t^k = 1)$ , and its value can be estimated through the time series data. The expected number of times node  $i$  is activated at time  $t_m + 1$  can be expressed as

$$\begin{aligned} E_i^{t_m+1} = & \sum_{j(j \neq i)} P_{j \rightarrow i} P_j^i \Psi_j^{t_m} \\ & + \sum_{j, k(j \neq k \neq i)} P_{jk \rightarrow i} P_{jk}^i \Psi_j^{t_m} \Psi_k^{t_m} + \varepsilon_i. \end{aligned} \quad (14)$$

Next, by assuming that the times of activations follows a Poisson distribution, the likelihood function can be written down and the EM algorithm is applied to solve the value of  $P_{jk \rightarrow i}$ , finally the 2-simplices in the simplicial complex are inferred. Of course, the 3-simplices or higher-order interactions in the simplicial complex can be inferred in a similar way. Recently, Liu *et al.* adopted vectorization expressions to improve the code running efficiency of the above method [71].

**Outlook.** — We have provided a mini-review regarding network reconstruction approaches from the observed data driven by the discrete state dynamics. These methods can reconstruct networks with known nodal dynamics or unknown nodal dynamics. What is more, some methods can be used to reconstruct single-layer networks, duplex networks, signed networks as well as simplicial complexes. Although significant progresses have been made in this field, there are still many unresolved issues worth further research:

Firstly, many researches have demonstrated that hypergraphs provide a flexible and natural modeling tool to model such complex relationships. For instance, in the basketball games, where each team's five players present on the court at each moment can form a hyperedge, or in a co-authorship network, where several authors collaborating on a paper can form a hyperedge. However, compared to the pairwise network or simplicial complexes, the structures of hypergraphs are more complicated. For example, a hypergraph can contain hyperedges with different sizes, a hyperedge with large size may contain some hyperedges with small size, or hypergraphs do not meet the constraint condition as simplicial complexes (*i.e.*, higher-order interactions should include lower-order interactions). Therefore, how to reconstruct hypergraphs from the discrete state data is a meaningful and challenging research area [72,73].

Secondly, existing network reconstruction approaches based on discrete state dynamical data usually require to sample very long time series to achieve high reconstruction accuracy [74]. Nevertheless, collecting such long time series in real-world scenarios is often difficult, and it also yields high computational complexity. One possible solution is to incorporate some *a priori* knowledge of the network structure into the proposed methods, such as heterogeneity [75], symmetry [18], etc., and they have been proven to be effective in reconstructing networks based on continuous state dynamics data. In addition, in practical situations, the structure of the network is often not entirely unknown. Therefore, it is possible to incorporate partial structural information with shorter time series data to address this dilemma.

Thirdly, for discrete state dynamics, existing reconstruction methods often assume that the states of nodes are two or at most three, such as SIS, Ising, voter, SIR models. In reality, many discrete state dynamics may have multiple states, such as opinion dynamics, where each



state can interact and transform each other. As a result, how to reconstruct networks for multi-state discrete state dynamics is a very challenging problem.

Fourthly, existing methods are difficult to reconstruct large-scale network. This difficulty is not only reflected in computational time, but also in the reconstruction accuracy. One possible way of reconstructing large-scale networks is to divide the large-scale network into blocks and then infer the structure of each block. Therefore, two key issues need to be studied: 1) how to block the large-scale network through the discrete dynamic data; 2) how to reconstruct a network with unclosed network structure (*i.e.*, the connections with other blocks are missed). If these two problems can be solved, this can lead the network reconstruction problem to scale and parallelize.

Fifthly, in many situations, some nodes' data may be fully unknown (*i.e.*, nodes are hidden) [76–80] or the size of the network is unknown [81]. Evidently, it is a quite challenging problem to reconstruct the network structure as well as to infer the existence and location of hidden nodes, only based on the available time series. Even some methods have been proposed to solve this problem, such as compressive sensing based method [76–78], random variable resetting based method [79], statistical inference based method [80], most existing methods were proposed for the continuous state dynamics data. So how to generalize the existing methods or develop special methods for the discrete state dynamics data case also requires focused research.

Lastly, most reconstruction methods based on discrete state dynamical data use the simulated dynamical data to reconstruct network structures, and the performance of these methods on real data is often unideal, such as the real data from disease or information flow. Thus, how to develop more effective and practical reconstruction approaches for real discrete data is a significant challenge. Addressing this challenge would greatly enhance the practical applications of network reconstruction in real-world scenarios.

\* \* \*

This work is partially supported by the National Natural Science Foundation of China (12005001, 61973001), and the University Synergy Innovation Program of Anhui Province (GXXT-2021-032).

*Data availability statement:* No new data were created or analysed in this study.

## REFERENCES

- [1] BOCCALETTI S., LATORA V., MORENO Y., CHAVEZ M. and HWANG D.-U., *Phys. Rep.*, **424** (2006) 175.
- [2] NEWMAN M. E., *SIAM Rev.*, **45** (2003) 167.
- [3] BOCCALETTI S., DE LELLIS P., DEL GENIO C., ALFARO-BITTNER K., CRIADO R., JALAN S. and ROMANCE M., *Phys. Rep.*, **1018** (2023) 1.
- [4] TIMME M. and CASADIEGO J., *J. Phys. A: Math. Theor.*, **47** (2014) 343001.
- [5] WANG W.-X., LAI Y.-C. and GREBOGI C., *Phys. Rep.*, **644** (2016) 1.
- [6] LI X. and LI X., *Nat. Commun.*, **8** (2017) 15729.
- [7] CASADIEGO J., NITZAN M., HALLERBERG S. and TIMME M., *Nat. Commun.*, **8** (2017) 2192.
- [8] CASADIEGO J., MAOUTSA D. and TIMME M., *Phys. Rev. Lett.*, **121** (2018) 054101.
- [9] WANG H., MA C., CHEN H.-S., LAI Y.-C. and ZHANG H.-F., *Nat. Commun.*, **13** (2022) 3043.
- [10] MA C., ZHANG H.-F. and LAI Y.-C., *Phys. Rev. E*, **96** (2017) 022320.
- [11] WANG W.-X., LAI Y.-C., GREBOGI C. and YE J., *Phys. Rev. X*, **1** (2011) 021021.
- [12] CHEN Y.-Z. and LAI Y.-C., *Phys. Rev. E*, **97** (2018) 032317.
- [13] WANG W.-X., YANG R., LAI Y.-C., KOVANIS V. and GREBOGI C., *Phys. Rev. Lett.*, **106** (2011) 154101.
- [14] SHEN Z., WANG W.-X., FAN Y., DI Z. and LAI Y.-C., *Nat. Commun.*, **5** (2014) 4323.
- [15] GAO T.-T. and YAN G., *Nat. Comput. Sci.*, **2** (2022) 160.
- [16] LEVNAJIC Z. and PIKOVSKY A., *Phys. Rev. Lett.*, **107** (2011) 034101.
- [17] MALIZIA F., CORSO A., GAMBUTTA L. V., RUSSO G., LATORA V. and FRASCA M., arXiv preprint, arXiv:2305.06609 (2023).
- [18] MA C., LAI Y.-C., LI X. and ZHANG H.-F., *Phys. Rev. E*, **108** (2023) 034304.
- [19] GARDNER T. S., DI BERNARDO D., LORENZ D. and COLLINS J. J., *Science*, **301** (2003) 102.
- [20] TEGNER J., YEUNG M. S., HASTY J. and COLLINS J. J., *Proc. Natl. Acad. Sci. U.S.A.*, **100** (2003) 5944.
- [21] YEUNG M. S., TEGNER J. and COLLINS J. J., *Proc. Natl. Acad. Sci. U.S.A.*, **99** (2002) 6163.
- [22] DI BERNARDO D., THOMPSON M. J., GARDNER T. S., CHOBOT S. E., EASTWOOD E. L., WOJTOVICH A. P., ELLIOTT S. J., SCHAUS S. E. and COLLINS J. J., *Nat. Biotechnol.*, **23** (2005) 377.
- [23] DE JONG H., *J. Comput. Biol.*, **9** (2002) 67.
- [24] LEE W.-P. and TZOU W.-S., *Briefings Bioinform.*, **10** (2009) 408.
- [25] REN J., WANG W.-X., LI B. and LAI Y.-C., *Phys. Rev. Lett.*, **104** (2010) 058701.
- [26] WANG W.-X., REN J., LAI Y.-C. and LI B., *Chaos*, **22** (2012) 033131.
- [27] ZHANG Z., CHEN Y., MI Y. and HU G., *Phys. Rev. E*, **99** (2019) 042311.
- [28] ZHANG Z., ZHENG Z., NIU H., MI Y., WU S. and HU G., *Phys. Rev. E*, **91** (2015) 012814.
- [29] TIMME M., *Phys. Rev. Lett.*, **98** (2007) 224101.
- [30] YU D., *Automatica*, **46** (2010) 2035.
- [31] YU D. and PARLITZ U., *Phys. Rev. E*, **82** (2010) 026108.
- [32] YU D., RIGHERO M. and KOCAREV L., *Phys. Rev. Lett.*, **97** (2006) 188701.
- [33] LIU H., LU J.-A., LÜ J. and HILL D. J., *Automatica*, **45** (2009) 1799.
- [34] WU X., ZHAO X., LÜ J., TANG L. and LU J.-A., *IEEE Trans. Control Netw. Syst.*, **3** (2015) 379.
- [35] ZHOU J. and LU J.-A., *Phys. A: Stat. Mech. Appl.*, **386** (2007) 481.

- [36] HAN X., SHEN Z., WANG W.-X. and DI Z., *Phys. Rev. Lett.*, **114** (2015) 028701.
- [37] SHANDILYA S. G. and TIMME M., *New J. Phys.*, **13** (2011) 013004.
- [38] HAN X., SHEN Z., WANG W.-X., LAI Y.-C. and GREBOGI C., *Sci. Rep.*, **6** (2016) 30241.
- [39] SRINIVASAN K., COBLE N., HAMLIN J., ANTONSEN T., OTT E. and GIRVAN M., *Phys. Rev. Lett.*, **128** (2022) 164101.
- [40] BANERJEE A., HART J. D., ROY R. and OTT E., *Phys. Rev. X*, **11** (2021) 031014.
- [41] LIU J., MEI G., WU X. and LÜ J., *IEEE Trans. Circuits Syst. I: Reg. Pap.*, **65** (2018) 2970.
- [42] ALDERISIO F., FIORE G. and DI BERNARDO M., *Phys. Rev. E*, **95** (2017) 042302.
- [43] LI G., LI N., LIU S. and WU X., *Chaos*, **29** (2019).
- [44] WU X., FAN Z., HE J., WANG W. and LÜ J., to be published in *IEEE Trans. Syst. Man Cybernet.: Syst.* (2023).
- [45] HU Z.-L., HAN X. and MA L., *IEEE Trans. Netw. Sci. Eng.*, **7** (2020) 3200.
- [46] PASTOR-SATORRAS R., CASTELLANO C., VAN MIEGHEM P. and VESPIGNANI A., *Rev. Mod. Phys.*, **87** (2015) 925.
- [47] SZABÓ G. and FATH G., *Phys. Rep.*, **446** (2007) 97.
- [48] CHEN X. and WANG L., *Phys. Rev. E*, **77** (2008) 017103.
- [49] SOOD V. and REDNER S., *Phys. Rev. Lett.*, **94** (2005) 178701.
- [50] DOROGOVTSSEV S. N., GOLTSEV A. V. and MENDES J. F. F., *Phys. Rev. E*, **66** (2002) 016104.
- [51] CANDES E. J. and TAO T., *IEEE Trans. Inf. Theory*, **52** (2006) 5406.
- [52] CANDÉS E. J., ROMBERG J. and TAO T., *IEEE Trans. Inf. Theory*, **52** (2006) 489.
- [53] MA C., CHEN H.-S., LI X., LAI Y.-C. and ZHANG H.-F., *SIAM J. Appl. Dyn. Syst.*, **19** (2020) 124.
- [54] LIU Q.-M., MA C., XIANG B.-B., CHEN H.-S. and ZHANG H.-F., *IEEE Trans. Syst. Man Cybernet.: Syst.*, **51** (2019) 4639.
- [55] XIANG B.-B., MA C., CHEN H.-S. and ZHANG H.-F., *Chaos*, **28** (2018).
- [56] GRAY C., MITCHELL L. and ROUGHAN M., *IEEE Trans. Signal Inf. Process. Netw.*, **6** (2020) 371.
- [57] BRAUNSTEIN A., INGROSSO A. and MUNTONI A. P., *J. R. Soc. Interface*, **16** (2019) 20180844.
- [58] LI J., SHEN Z., WANG W.-X., GREBOGI C. and LAI Y.-C., *Phys. Rev. E*, **95** (2017) 032303.
- [59] MA C., CHEN H.-S., LAI Y.-C. and ZHANG H.-F., *Phys. Rev. E*, **97** (2018) 022301.
- [60] LI X., LIN L., LIU Q., KONG Z. and DING L., *IEEE Trans. Comput. Soc. Syst.*, **10** (2021) 24.
- [61] ZHANG H.-F., XU F., BAO Z.-K. and MA C., *IEEE Trans. Circuits Syst. I: Reg. Pap.*, **66** (2018) 1608.
- [62] ZHANG Y., GUO Y., ZHANG Z., CHEN M., WANG S. and ZHANG J., *Phys. Rev. E*, **106** (2022) 034315.
- [63] GOETZE F. and LAI P.-Y., *Phys. Rev. E*, **100** (2019) 012121.
- [64] LAMBIOTTE R., ROSVALL M. and SCHOLTES I., *Nat. Phys.*, **15** (2019) 313.
- [65] SANTORO A., BATTISTON F., PETRI G. and AMICO E., *Nat. Phys.*, **19** (2023) 221.
- [66] IACOPINI I., PETRI G., BARRAT A. and LATORA V., *Nat. Commun.*, **10** (2019) 2485.
- [67] WANG H., ZHANG H.-F., ZHU P.-C. and MA C., *Chaos*, **32** (2022).
- [68] COURTNEY O. T. and BIANCONI G., *Phys. Rev. E*, **93** (2016) 062311.
- [69] PETRI G. and BARRAT A., *Phys. Rev. Lett.*, **121** (2018) 228301.
- [70] SHI D., CHEN Z., MA C. and CHEN G., *Research*, **6** (2023) 0230.
- [71] LIU K., LÜ X., GAO F. and ZHANG J., *Phys. D: Nonlinear Phenom.*, **454** (2023) 133834.
- [72] LIZOTTE S., YOUNG J.-G. and ALLARD A., arXiv preprint, arXiv:2208.06503 (2022).
- [73] MARINAZZO D., VAN ROOZENDAAL J., ROSAS F. E., STELLA M., COMOLATTI R., COLENBIER N., STRAMAGLIA S. and ROSSEEL Y., arXiv preprint, arXiv:2205.01035 (2022).
- [74] PRASSE B. and VAN MIEGHEM P., *IEEE Trans. Netw. Sci. Eng.*, **6** (2018) 748.
- [75] MA L., HAN X., SHEN Z., WANG W.-X. and DI Z., *PLoS ONE*, **10** (2015) e0142837.
- [76] SU R.-Q., LAI Y.-C., WANG X. and DO Y., *Sci. Rep.*, **4** (2014) 3944.
- [77] SU R.-Q., WANG W.-X., WANG X. and LAI Y.-C., *R. Soc. Open Sci.*, **3** (2016) 150577.
- [78] SU R.-Q., WANG W.-X. and LAI Y.-C., *Phys. Rev. E*, **85** (2012) 065201.
- [79] JIANG W. and WANG S., *Chaos*, **33** (2023) 043109.
- [80] ZHANG Y., LI H., ZHANG Z., QIAN Y. and PANDEY V., *Chin. J. Phys.*, **67** (2020) 203.
- [81] HAEHNE H., CASADIEGO J., PEINKE J. and TIMME M., *Phys. Rev. Lett.*, **122** (2019) 158301.