

Data based reconstruction of complex multiplex networks

Chuang Ma,¹ Han-Shuang Chen,² Xiang Li,³ Ying-Cheng Lai,⁴ and Hai-Feng Zhang^{1,5,*}

¹*School of Mathematical Science, Anhui University, Hefei 230601, China*

²*School of Physics and Material Science, Anhui University, Hefei 230601, China*

³*Adaptive Networks and Control Laboratory, Department of Electronic Engineering,
and the Center of Smart Networks and Systems,*

School of Information Science and Engineering, Fudan University, Shanghai 200433, China

⁴*School of Electrical, Computer and Energy Engineering,*

Arizona State University, Tempe, Arizona 85287, USA

⁵*Department of Communication Engineering, North University of China, Taiyuan, Shan'xi 030051, China*

(Dated: August 29, 2018)

It has been recognized that many complex dynamical systems in the real world require a description in terms of multiplex networks, where a set of common, mutually connected nodes belong to distinct network layers and play a different role in each layer. In spite of recent progress towards data based inference of single-layer networks, to reconstruct complex systems with a multiplex structure remains largely open. We articulate a mean-field based maximum likelihood estimation framework to solve this outstanding and challenging problem. We demonstrate the power of the reconstruction framework and characterize its performance using binary time series from a class of prototypical duplex network systems that host two distinct types of spreading dynamics. In addition to validating the framework using synthetic and real-world multiplex networks, we carry out a detailed analysis to elucidate the impacts of structural and dynamical parameters as well as noise on the reconstruction accuracy and robustness.

INTRODUCTION

In physical and mathematical sciences, it is recognized that the “inverse problem” is often significantly more difficult than the “forward problem.” In particular, given a system with a known structure and a set of mathematical equations, the forward problem focuses on analyzing and possibly solving the equations (analytically or numerically) to uncover and understand the behaviors of the system. For the inverse problem, the system structure and equations are unknown but only observational or measured data are available. The task is to infer the intrinsic structure and dynamics of the system from the data. In network science and engineering, to reconstruct the topology of an unknown complex network and to map out the dynamical process on the network based solely on measured time series or data have been an active area of interdisciplinary research [1–31]. A variety of approaches have been devised, which include those based on collective dynamics [6, 8, 21, 23, 32], stochastic analysis [29, 33], and compressive sensing [24, 25, 31, 34], etc. However, previous works focused on single-layer networks. The goal of this paper is to address the significantly more challenging problem of data based reconstruction of *multiplex* networks.

A complex system in the real world, such as the modern infrastructure, a social or a transportation system, consists of many units connected by different types of relationship. For example, a social network contains different types of ties among people and a transportation system comprises multiple types of travel platforms. Such systems require a description in terms of multiplex networks [35–39]. Previous efforts in multiplex networks focused on the forward problem to unearth the mathematical properties and the associated physical phenomena. The main difficulty that one has to overcome to address the inverse problem of multiplex networked systems lies in the distinct, possibly quite diverse yet interwoven collective dynamics in different layers. For example, the outbreak of an epidemic in the human society induces diffusion of awareness in online social networks, leading to two types of mutually coupled spreading dynamics [40], each lives in a different network layer. Another example is that opinions can diffuse through different channels (layers) and interact with each other. To our knowledge, there has been no prior work on the inverse problem for multiplex networks.

In this paper, we develop a reconstruction framework based on mean-field maximum likelihood estimation to address the problem of data based reconstruction of multiplex networks. To describe our framework in a concrete and clear way, we focus on duplex networked systems - perhaps the most extensively studied multiplex networks that are relevant to real world situations such as complex cyberphysical systems. We assume that each layer hosts a distinct type of spreading dynamics and the two types of processes are interwoven. In particular, one (physical) layer hosts the susceptible-infected-susceptible (SIS) type of spreading dynamics, while the other (virtual) layer is a social network with information spreading governed by the unaware-aware-unaware (UAU) process [40]. Provided that binary time series data are available from both layers, we show that our framework is capable of accurately reconstructing the full topology of each layer for a large number of empirical and synthetic networks. We elucidate the impacts of network structural and dynamics parameters on the reconstruction accuracy, such as the average degree, interlayer coupling, and heterogeneity in the spreading rates. The effect of noise is also investigated. Our framework represents a success to assess the “internal gear” of complex systems with a duplex structure, which can be generalized to networks with a more sophisticated structure and diverse dynamical processes.

RESULTS

Basic notions of reconstruction framework. Duplex networks with UAU-SIS coupled dynamics describe realistic situations where there is competition between disease spreading and social awareness, with the physical contact layer supporting an epidemic process and the virtual contact layer generating awareness diffusion, where the former is of the SIS type while the latter can be described by the UAU process.

The mutual interplay or coupling between the two distinct types of spreading processes can be described, as follows. An I-node in the physical layer is automatically aware of the infection and changes to the A (aware) state immediately in the virtual layer. If an S-node is in the A state, its ability to infect other connected nodes is discounted by a factor $0 \leq \gamma < 1$. For convenience, we introduce the mathematical notation ${}_{x_4}^{x_3} X {}_{x_1}^{x_2}$ associated with a variable X , where x_1 and x_2 are the time attribute and the id number of X (e.g., $x_1 = t$ and $x_2 = i$), x_3 determines whether X belongs to the virtual layer (i.e., $x_3 = 1$) or the physical contact layer (i.e., $x_3 = 2$), x_4 specifies either U or A. For example, ${}_U\beta$ and ${}_A\beta = \gamma \cdot {}_U\beta$ denote the infection rate of an unaware and an aware S-node, respectively. Variable s with three annotations denotes the state of a node, e.g., ${}^1s_{t_m}^i = 0$ or 1 (${}^2s_{t_m}^i = 0$ or 1) indicates that node i in the virtual (physical) layer is in U or A state (S or I state) at time t_m . The framework can deal with the general case where the transmission and/or recovery rates are heterogeneous among the nodes.

Our reconstruction framework consists of three steps: (a) to formulate the likelihood function of the coupled dynamics, (b) to perform the mean-field based approximation to enable the estimation, and (c) to transform the estimation problem into two solvable linear systems - one for each layer with solutions representing the neighbors

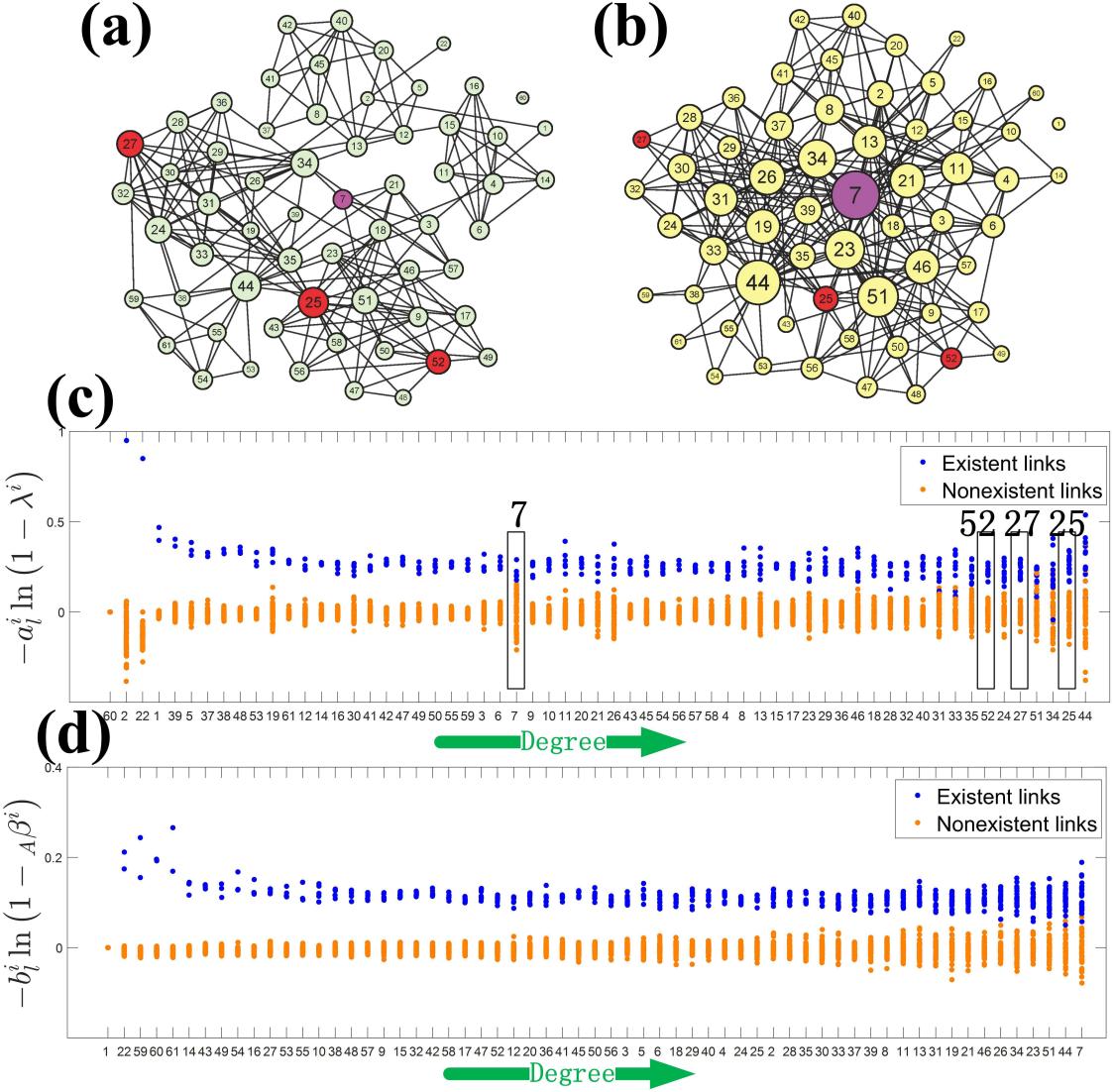


FIG. 1. Reconstruction of the CS-AARHUS network. (a) Actual structure of the virtual contact layer (Facebook). (b) The structure of the physical layer. (c,d) The values of $-a_l^i \ln(1 - \lambda^i)$, $i \neq l$ and $-b_l^i \ln(1 - \lambda^i)$, $i \neq l$, respectively, versus the nodal degree. Each column gives the connectivity of a node, where the blue and orange dots denote the existent and nonexistent links, respectively. The length of the time series is $M = 30000$. The parameter values for the dynamical processes are $\lambda = 0.2$, $U\beta = 0.2$, $A\beta = 0.5U\beta$, and $\mu = \delta = 0.8$.

of each node in the layer. A detailed formulation of the reconstruction framework is presented in **Methods** and a schematic illustration of the UAU-SIS coupling dynamics is given in Fig. 6 of Supplementary Information (SI).

Performance demonstration: reconstruction of three real-world duplex networks. For a network layer (say the virtual layer), the mathematical formulation of our reconstruction framework gives a characteristic, node-specific quantity for distinguishing the existent from the non-existent links: $-a_l^i \ln(1 - \lambda^i)$, where \mathbf{a}^i is the vector defining the local connection structure of node i : $a_l^i = 1$ specifies that node l is in the neighborhood of i and $a_j^i = 0$ otherwise, and λ^i is the probability that node i (if it is unaware) is informed by an aware (A) neighbor. For actual links, the values of $-a_l^i \ln(1 - \lambda^i)$ are finite and away from zero, while the values for non-existent links are near zero. If there is a gap between the two sets of values, the actual links can be detected. A similar quantity can be defined for the physical layer, which is denoted as $-b_l^i \ln(1 - A\beta^i)$.

We first validate our framework using an empirical network of 61 nodes, the so-called CS-AARHUS network [41]. The original network has five layers. We regard the Facebook layer as the virtual layer and the other four off-line layers (Leisure, Work, Co-authorship, Lunch) as the physical layer, as illustrated in Figs. 1(a,b), respectively. Figures 1(c,d) show the values of characteristic quantities $-a_l^i \ln(1 - \lambda^i)$ and $-b_l^i \ln(1 - A\beta^i)$ for the virtual and physical layers,

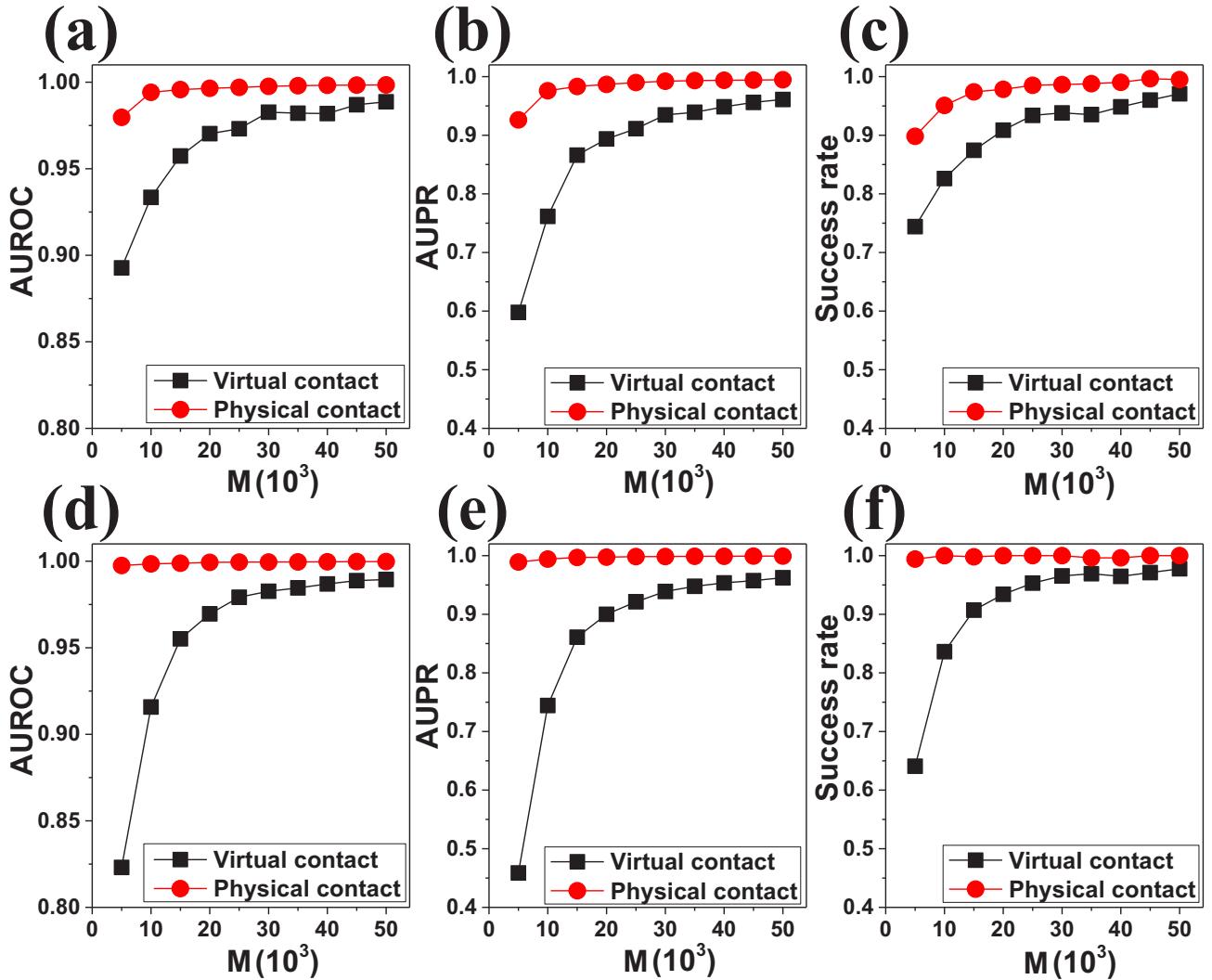


FIG. 2. **Reconstruction accuracy of *C. elegans* and CKM networks.** (a-c) For the *C. elegans* network, values of AUROC, AUPR and Success rate versus the length M of the time series, respectively, for the parameter setting $\lambda = 0.2$, $U\beta = 0.3$, $A\beta = 0.5U\beta$, and $\mu = \delta = 0.8$. (d-f) The corresponding results for the CKM network for the same parameter values as for (a-c) except $U\beta = 0.5$.

where the blue and orange dots denote the actual and nonexistent links, respectively. We see that the values of the characteristic quantities are well separated by a distinct gap and can be unequivocally distinguished through a properly chosen threshold (see also Fig. 8 in SI). For the physical layer in Fig. 1(d), the gap between the blue and orange dots exhibits a decreasing trend with the nodal degree, indicating that the neighbors of larger degree nodes are harder to be detected because of neighborhood overlapping associated with such nodes. This result is consistent with previous findings [25, 33]. For the virtual layer [Fig. 1(c)], the blue and orange dots for node 7 are overlapped even though ${}^1k^7 = 6$, but there is a finite gap for large degree nodes, e.g., node 52 with ${}^1k^{52} = 10$, node 27 with ${}^1k^{27} = 12$, and node 25 with ${}^1k^{25} = 15$. The relatively small gap of ${}^1k^7$ is due to the fact that the counterpart value in the physical layer is large: ${}^2k^7 = 29$, indicating that the node has been infected and is thus constantly in the A state in the virtual layer (an infected node becomes aware immediately). As a result, the states of the neighbors of this node in the virtual layer have little influence on its state, making reconstruction difficult. For nodes with large and small degrees in the virtual and physical layers, respectively, the transition from U to A is mainly determined by the states of the neighbors, facilitating reconstruction. In general, the structure of the physical layer has a significant effect on the reconstruction of the virtual layer, but the effect in the opposite direction is minimal.

We next demonstrate the power of our reconstruction framework for two additional empirical networks: *C. elegans* [39] and the innovation diffusion network for physicians (the so-called CKM network) [42]. Each network has three layers. We choose the first and the second layer as the virtual and physical layer, respectively. The panels in

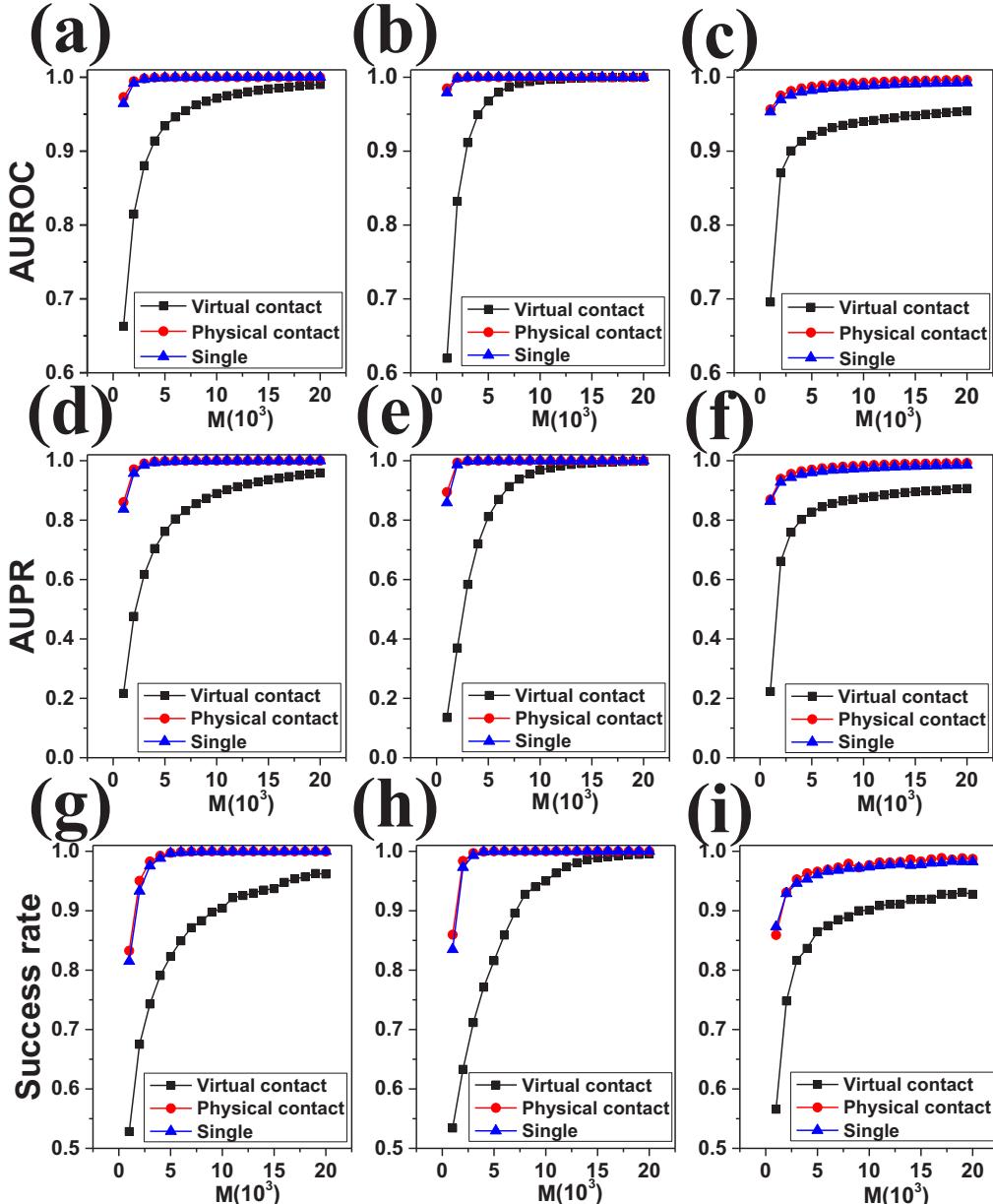


FIG. 3. **Effect of interlayer coupling on reconstruction accuracy.** Columns 1-3: reconstruction performance for ER-ER, SW-SW, and BA-BA duplex networks, respectively. The “single” case indicates the absence of interlayer coupling: $u\beta = 0$ ($\lambda = 0$) for the virtual (physical) layer. The parameter setting is $\lambda = 0.4$, $u\beta = 0.4$, $A\beta = 0.5u\beta$, and $\mu = \delta = 0.8$. The structures of the two layers are identical. The network parameters are $N = 100$ and $\langle k_1 \rangle = \langle k_2 \rangle = 6$.

the top and bottom rows of Fig. 2 display the reconstruction accuracy in terms of the statistical quantities AUROC, AUPR and Success rate (**Methods**) versus the length of the time series for the *C. elegans* and CKM networks, respectively. We have that longer time series result in better reconstruction performance and the reconstruction accuracy of the physical layer is higher than that of the virtual layer, which are consistent with the results in Fig. 1.

Performance analysis: reconstruction of synthetic multiplex networks. To understand the effect of interlayer coupling on reconstruction, we test a number of synthetic complex networks: small-world (SW-SW) [43], Erdős-Rényi (ER-ER) [44], and Barabási-Albert (BA-BA) [45] duplex systems. For comparison, we include the special case where each layer is separately reconstructed without taking into account the other layer, which is equivalent to reconstructing a single-layer network (labeled as single). Figures 3(a-i) show that the reconstruction accuracy of the virtual layer is greatly reduced when a physical layer is introduced (e.g., blue/red → black). Without the physical layer,

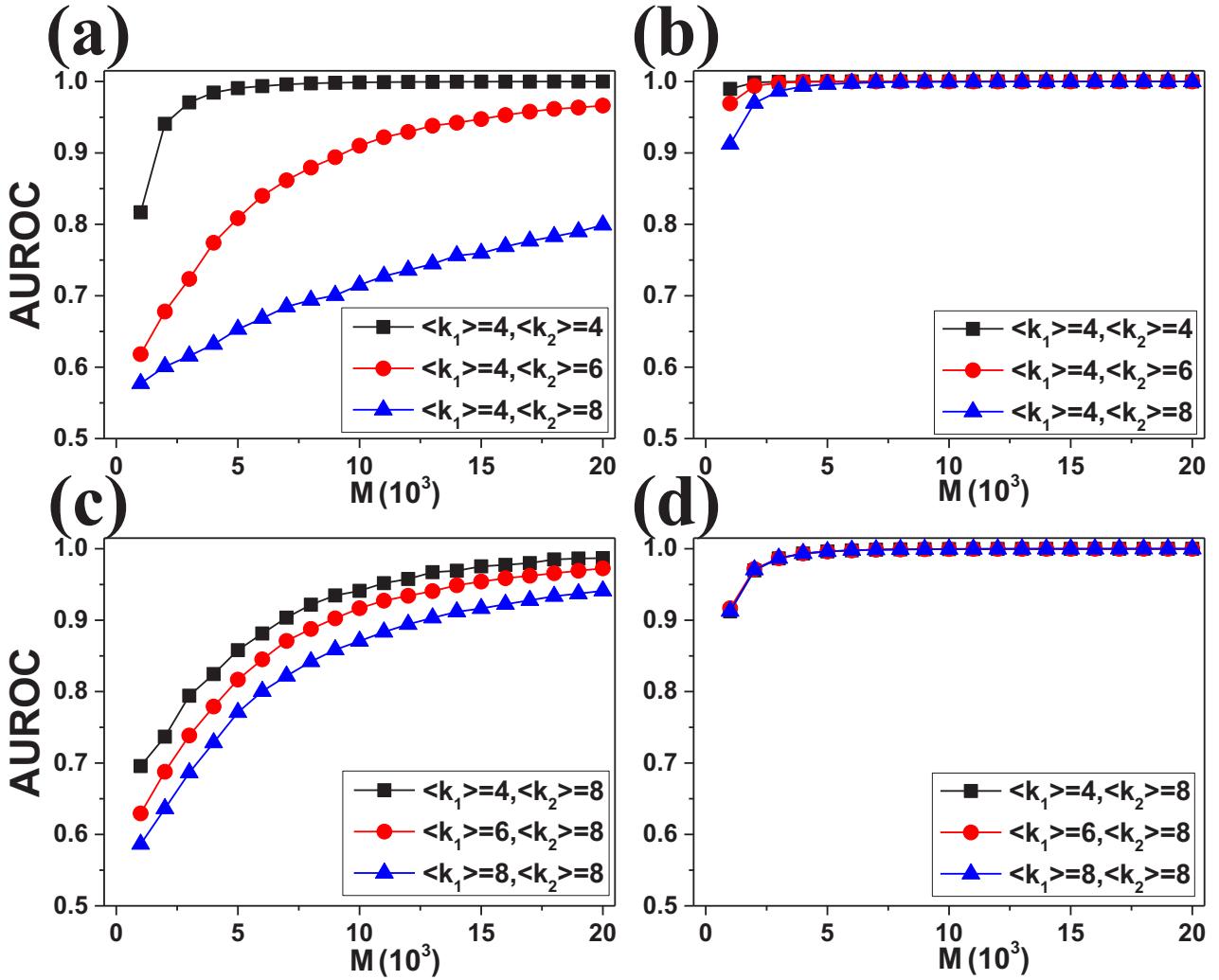


FIG. 4. Effect of average degree on reconstruction as measured by the AUROC index. (a,b) For a fixed value of the average degree $\langle k_1 \rangle$ of the virtual layer, the effect of varying the average degree $\langle k_2 \rangle$ of the physical layer on the reconstruction accuracy of the former and latter, respectively. (c,d) For a fixed value of $\langle k_2 \rangle$, the effect of varying the value of $\langle k_1 \rangle$ on the reconstruction accuracy of the virtual and physical layer, respectively. ER-ER duplex networks with $N=100$ are used. The parameters are $\lambda = 0.3$, $U\beta = 0.4$, $A\beta = 0.5U\beta$, and $\mu = \delta = 0.8$.

the transition of an unaware node in the virtual layer to the aware state depends only on the states of its neighbors. With the presence of the physical layer, an A-node can spontaneously become aware once it is infected, “concealing” the information about the structure of the virtual layer. On the contrary, the reconstruction accuracy of the physical layer can be improved slightly (e.g., blue versus red) when the virtual layer is introduced, which reduces the ability to infect of A-nodes and prevents too many nodes from being in the I state, facilitating reconstruction. Figure 3 also illustrates that the reconstruction accuracy of the SW-SW duplex network is higher than that of the ER-ER duplex network and much higher than that of the BA-BA duplex network due to the difficulty in reconstructing the neighbors of large degree nodes.

How does the average degree of each layer affect the reconstruction accuracy? Figure 4(a) shows that an increase in the average degree $\langle k_2 \rangle$ of the physical layer can greatly reduce the reconstruction accuracy of the virtual layer. Figures 4(b,c) show that, for the physical layer, the accuracy gradually decreases with its average degree, for a fixed average degree of the virtual layer. An explanation is that the probability of being infected tends to increase for a larger value of $\langle k_1 \rangle$, “hiding” the information required for uncovering the structure of the virtual layer. We also find that increasing the average degree $\langle k_1 \rangle$ of the virtual layer tends to reduce the reconstruction accuracy of itself [Fig. 4(c)] but has a negligible effect on the reconstruction of the physical layer [Fig. 4(d)].

Figure 5 shows the effect of noise on the reconstruction accuracy, where noise is implemented by randomly flipping a fraction τ of the states among the total number MN of states. Noise has a significant effect on the reconstruction

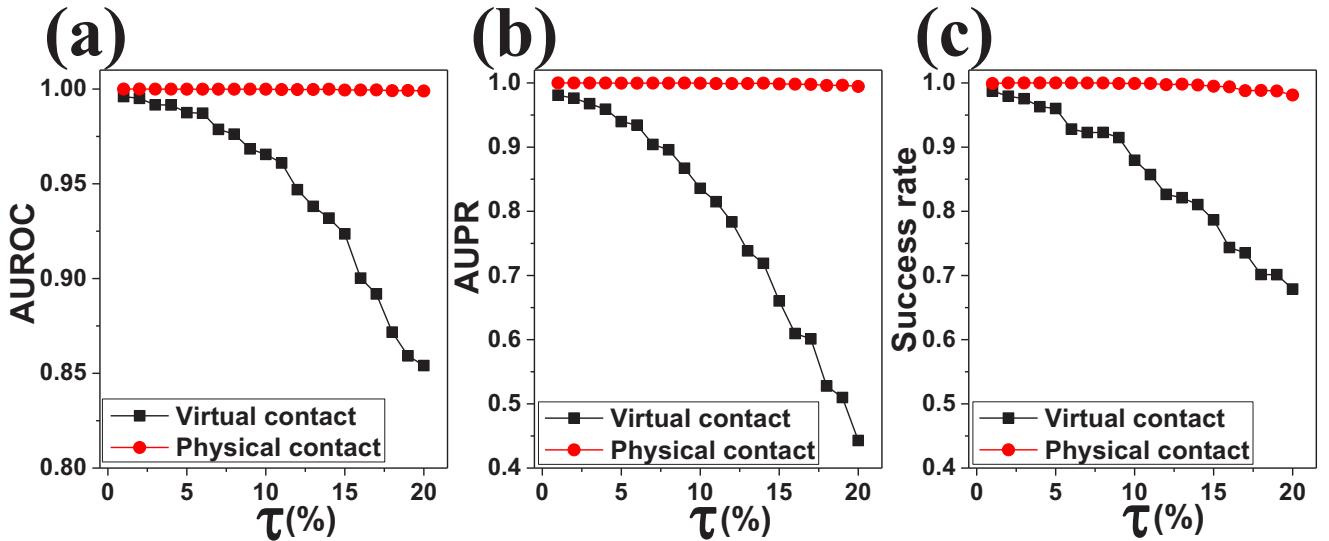


FIG. 5. **Impact of noise on reconstruction accuracy.** (a-c) AUROC, AUPR and Success rate versus the fraction τ of randomly flipped states for an ER-ER duplex system. The network parameters are $N = 100$, $\langle k_1 \rangle = 4$, and $\langle k_2 \rangle = 6$. The length of the time series is $M = 30000$. Other parameters are the same as in Fig. 4.

of the virtual layer, but it hardly affects the reconstruction of the physical layer (even when the flip rate is $\tau = 20\%$).

DISCUSSION

We have developed a mean-field based maximum likelihood estimation framework to solve the challenging problem of data based reconstruction of multiplex networks. The reconstruction performance has been demonstrated using a number of real-world and synthetic duplex networks comprising a virtual and a physical layer, where each layer hosts a distinct type of spreading dynamics that are coupled through the duplex network structure. Extensive test and analysis indicate that the framework is capable of accurately reconstructing the full topology of each layer based solely on measured time series. A thorough examination of the dynamical coupling between the two layers gives that the reconstruction accuracy of the physical layer is generally much higher than that of the virtual layer. In addition, the reconstruction accuracy of the virtual layer is more sensitive to external noise than that of the physical layer.

Our framework represents a starting point towards reconstructing more general multiplex networks hosting different types of dynamics. Appealing features are that the framework is parameter free, has high accuracy, is readily to be implemented, and has a solid mathematical foundation. Issues warranting further considerations include extension to continuous-time dynamical processes, generalization to multiplex networks consisting of more than two layers, and development of effective and practical methods to reduce the required data amount.

METHODS

UAU-SIS dynamics on duplex networks. The UAU-SIS model was originally articulated to study the competition between social awareness and disease spreading on double layer networks, where the physical contact layer supports an epidemic process and the virtual contact layer supports awareness diffusion [40]. The two layers share exactly the same set of nodes but their connection patterns are different. Spreading of awareness in the virtual layer is described by the classic SIS type of epidemic dynamics: an unaware node (U) is informed by an aware neighbor (A) with probability λ , and an aware node can lose awareness and returns to the U state with probability δ . Epidemic dynamics in the physical layer are also of the SIS type, where an infected (I) node can infect its susceptible (S) neighbors with probability β , and an I-node returns to the S state with probability μ . The interlayer interaction is described, as follows. An I-node in the physical layer is automatically aware of the infection and changes to the A state immediately in the virtual layer. If an S-node is in the A state, its ability to infect is reduced by a factor $0 \leq \gamma < 1$. Figure 6 in SI presents a schematic illustration of the duplex network with the described interacting dynamical process.

Mathematical formulation of the reconstruction framework. For a duplex network hosting UAU-SIS dynamics, at any time a node can be in one of the three distinct states: unaware and susceptible (US), aware and susceptible (AS), or aware and infected (AI). The state UI (unaware and infected) is redundant because an I-node becomes aware immediately. The connections of node i in the virtual and physical layers are specified by the vectors \mathbf{a}^i and \mathbf{b}^i , respectively, where $a_j^i = 1$ indicates that node j is a neighbor of node i in the virtual layer and $a_j^i = 0$ otherwise, and b_j^i is defined similarly.

The various transition probabilities can be derived from the transition tree of the UAU-SIS model (Fig. 7 in SI). Let ${}_U\beta$ and ${}_A\beta = \gamma \cdot {}_U\beta$ be the infection rates of unaware and aware S-node, respectively. For node i , $\sum_{j \neq i} a_j^{i-1} s_t^j$ and $\sum_{j \neq i} b_j^{i-2} s_t^j$ are the numbers of A-neighbors in the virtual layer and I-neighbors in the physical layer, respectively. The following probabilities are essential to the spreading dynamics: (1) r_t^i , the probability that node i is not informed by any neighbor, (2) ${}_U q_t^i$, the probability that node i is not infected by any neighbor if i was unaware, and (3) ${}_A q_t^i$, the probability that node i is not infected by any neighbor if i was aware. In the absence of any dynamical correlation, these probabilities are given by

$$\begin{aligned} r_t^i &= (1 - \lambda^i)^{\sum_{j \neq i} a_j^{i-1} s_t^j}, \\ {}_U q_t^i &= (1 - {}_U \beta^i)^{\sum_{j \neq i} b_j^{i-2} s_t^j}, \\ {}_A q_t^i &= (1 - {}_A \beta^i)^{\sum_{j \neq i} b_j^{i-2} s_t^j}. \end{aligned} \quad (1)$$

A tacit assumption in Ref. [40] is that diffusion of awareness in the virtual layer occurs before epidemic spreading in the physical layer. In our work, we do not require that the two types of spreading dynamics occur in any particular order.

From Fig. 7 in SI and Eq. (1), we have that the probabilities of node i being in the US, AS and AI states at $t + 1$ when it is in the US state at time t are

$$\begin{aligned} P^{US \rightarrow US} &= r_t^i {}_U q_t^i, \\ P^{US \rightarrow AS} &= (1 - r_t^i) {}_U q_t^i, \\ P^{US \rightarrow AI} &= r_t^i (1 - {}_U q_t^i) + (1 - r_t^i) (1 - {}_U q_t^i) = 1 - {}_U q_t^i. \end{aligned} \quad (2)$$

The probabilities of node i being in the US, AS and AI states at $t + 1$ when it is in the AS state at time t are

$$\begin{aligned} P^{AS \rightarrow US} &= \delta^i {}_A q_t^i, \\ P^{AS \rightarrow AS} &= (1 - \delta^i) {}_A q_t^i, \\ P^{AS \rightarrow AI} &= \delta^i (1 - {}_A q_t^i) + (1 - \delta^i) (1 - {}_A q_t^i) = 1 - {}_A q_t^i. \end{aligned} \quad (3)$$

The probabilities for node i to be in the US, AS and AI states at $t + 1$ if it is in the AI state at time t are

$$\begin{aligned} P^{AI \rightarrow US} &= \delta^i \mu^i, \\ P^{AI \rightarrow AS} &= (1 - \delta^i) \mu^i, \\ P^{AI \rightarrow AI} &= 1 - \mu^i. \end{aligned} \quad (4)$$

Say only the states ${}^1 s_{t_m+1}^i$ and ${}^2 s_{t_m+1}^i$ ($i = 1, \dots, N$) at time t_m (not necessarily uniform) are recorded, where N is the network size. Our framework consists of three steps: (1) to reconstruct the likelihood function of the coupled dynamics, (2) to apply the mean-field approximation to enable maximum likelihood estimation (MLE), and (3) to transform the MLE problem into two solvable linear systems - one for each layer with solutions representing the neighbors of each node in the layer. In particular, the likelihood function of node i is given by (see Sec. III of SI):

$$\begin{aligned} &P \left(\left\{ {}^1 s_{t_m+1}^i, {}^2 s_{t_m+1}^i \right\}_{m=1 \dots M} \mid \left\{ {}^1 s_{t_m}^j, {}^2 s_{t_m}^j \right\}_{j=1 \dots N, m=1 \dots M}, \mathbf{a}^i, \mathbf{b}^i, \lambda^i, {}_U \beta^i, {}_A \beta^i, \delta^i, \mu^i \right) \\ &= \prod_m \left[(r_{t_m}^i {}_U q_{t_m}^i)^{\left(1 - {}^1 s_{t_m+1}^i\right)\left(1 - {}^2 s_{t_m+1}^i\right)} ((1 - r_{t_m}^i) {}_U q_{t_m}^i)^{\left(1 - {}^1 s_{t_m+1}^i\right)\left(1 - {}^2 s_{t_m+1}^i\right)} (1 - {}_U q_{t_m}^i)^{\left(1 - {}^1 s_{t_m+1}^i\right)\left(1 - {}^2 s_{t_m+1}^i\right)} \right]^{\left(1 - {}^1 s_{t_m}^i\right)\left(1 - {}^2 s_{t_m}^i\right)} \\ &\quad \left[(\delta^i {}_A q_{t_m}^i)^{\left(1 - {}^1 s_{t_m+1}^i\right)\left(1 - {}^2 s_{t_m+1}^i\right)} ((1 - \delta^i) {}_A q_{t_m}^i)^{\left(1 - {}^1 s_{t_m+1}^i\right)\left(1 - {}^2 s_{t_m+1}^i\right)} (1 - {}_A q_{t_m}^i)^{\left(1 - {}^1 s_{t_m+1}^i\right)\left(1 - {}^2 s_{t_m+1}^i\right)} \right]^{\left(1 - {}^1 s_{t_m}^i\right)\left(1 - {}^2 s_{t_m}^i\right)} \\ &\quad \left[(\delta^i \mu^i)^{\left(1 - {}^1 s_{t_m+1}^i\right)\left(1 - {}^2 s_{t_m+1}^i\right)} ((1 - \delta^i) \mu^i)^{\left(1 - {}^1 s_{t_m+1}^i\right)\left(1 - {}^2 s_{t_m+1}^i\right)} (1 - \mu^i)^{\left(1 - {}^1 s_{t_m+1}^i\right)\left(1 - {}^2 s_{t_m+1}^i\right)} \right]^{\left(1 - {}^1 s_{t_m}^i\right)\left(1 - {}^2 s_{t_m}^i\right)}, \end{aligned} \quad (5)$$

where M is the length of the time series (step 1). The logarithmic form of Eq. (5), denoted as

$$L = L(a^i, b^i, \lambda^i, {}_U\beta^i, {}_A\beta^i, \delta^i, \mu^i),$$

can be written as

$$L = L_0(\delta^i, \mu^i) + L_1(\mathbf{a}^i, \lambda^i) + L_2(\mathbf{b}^i, {}_U\beta^i, {}_A\beta^i).$$

MLE can be performed by maximizing L_0 , L_1 and L_2 (Sec. III of SI). Because L_0 does not contain information about the network structure, it is only necessary to maximize L_1 and L_2 with respect to a_j^i and b_j^i , respectively. The likelihood function for reconstructing the virtual layer is (see Sec. III of SI for the corresponding function for the physical layer):

$$L_1(\mathbf{a}^i, \lambda^i) = \sum_m {}^1X_{t_m}^i \ln \left((1 - \lambda^i)^{\sum_{j \neq i} a_j^i {}^1s_{t_m}^j} \right) + \sum_m {}^1Y_{t_m}^i \ln \left(1 - (1 - \lambda^i)^{\sum_{j \neq i} a_j^i {}^1s_{t_m}^j} \right) \quad (6)$$

where

$$\begin{aligned} {}^1X_{t_m}^i &= (1 - {}^1s_{t_m}^i)(1 - {}^1s_{t_m+1}^i) \text{ and} \\ {}^1Y_{t_m}^i &= (1 - {}^1s_{t_m}^i)(1 - {}^2s_{t_m+1}^i){}^1s_{t_m+1}^i. \end{aligned}$$

The maximum value of L_1 cannot be obtained straightforwardly by setting zero its derivative with respect to a_j^i , because a_j^i appears in the exponential term and the values of λ^i are unknown. We resort to the mean-field approximation to solve this problem (step 2). Specifically, for node i in the virtual layer, the fraction $\sum_{j \neq i} {}^1s_{t_m}^j a_j^i$ of A-neighbors is approximately equal to the fraction of A-nodes in the whole layer excluding node i itself:

$$\sum_{j \neq i} {}^1s_{t_m}^j a_j^i \approx \frac{{}^1k^i}{N-1} {}^1\theta_{t_m}^i, \quad (7)$$

where ${}^1k^i$ is the degree of node i , ${}^1\theta_{t_m}^i = \sum_{j \neq i} {}^1s_{t_m}^j$ is the number of A-nodes excluding node i itself. Letting

$${}^1\gamma^i = (1 - \lambda^i)^{\frac{{}^1k^i}{N-1}},$$

we rewrite Eq. (6) concisely as

$$\hat{L}_1({}^1\gamma^i) = \sum_m \left[{}^1X_{t_m}^i \ln \left(({}^1\gamma^i)^{{}^1\theta_{t_m}^i} \right) + {}^1Y_{t_m}^i \ln \left(1 - ({}^1\gamma^i)^{{}^1\theta_{t_m}^i} \right) \right]. \quad (8)$$

Differentiating $\hat{L}_1({}^1\gamma^i)$ with respect to ${}^1\gamma^i$ and setting it to zero, we get the equation

$$\sum_m {}^1Y_{t_m}^i {}^1\theta_{t_m}^i \frac{({}^1\gamma^i)^{{}^1\theta_{t_m}^i}}{1 - ({}^1\gamma^i)^{{}^1\theta_{t_m}^i}} = \sum_m {}^1X_{t_m}^i {}^1\theta_{t_m}^i, \quad (9)$$

whose solutions give the values of ${}^1\gamma^i$ denoted as ${}^1\gamma^i = {}^1\tilde{\gamma}^i$.

To transform the MLE problem into a linear system (step 3), we take the derivative of Eq. (6) with respect to a_l^i and set it to zero, which gives the following high-dimensional nonlinear equation (the detailed reconstruction process for the virtual and physical layers is given in Secs. IV and V of SI, respectively):

$$\sum_m {}^1Y_{t_m}^i {}^1s_{t_m}^l \frac{(1 - \lambda^i)^{\sum_{j \neq i} a_j^i {}^1s_{t_m}^j}}{1 - (1 - \lambda^i)^{\sum_{j \neq i} a_j^i {}^1s_{t_m}^j}} = \sum_m {}^1X_{t_m}^i {}^1s_{t_m}^l.$$

To obtain a solution, we take advantage of Eq. (7) and use the first-order approximation. The result is a solvable linear system expressed in the matrix form

$$\boldsymbol{\Lambda} \cdot \boldsymbol{\eta} = \boldsymbol{\zeta}, \quad (10)$$

with the vector

$$\boldsymbol{\eta} = [a_1^i \ln(1 - \lambda^i), \dots, a_{i-1}^i \ln(1 - \lambda^i), a_{i+1}^i \ln(1 - \lambda^i), \dots, a_N^i \ln(1 - \lambda^i)]^T$$

(see Sec. IV of SI for the detailed forms of $\boldsymbol{\Lambda}$ and $\boldsymbol{\zeta}$). With the available time series ${}^1s_{t_m}^i$ and ${}^2s_{t_m}^i$, the values of the elements of the matrix $\boldsymbol{\Lambda}$ and the components of the vector $\boldsymbol{\zeta}$ can be calculated, so that the vector $\boldsymbol{\eta}$ characterizing the connectivity of node i can be solved. Because $\ln(1 - \lambda^i) < 0$ is a constant, the value of $-a_l^i \ln(1 - \lambda^i)$ is much above zero for $a_l^i = 1$ or close to zero for $a_l^i = 0$. A threshold value can then be readily set to distinguish the existent from the nonexistent links: a pair of nodes i and l are connected in the virtual layer if the value of $-a_l^i \ln(1 - \lambda^i)$ is larger than the threshold (see Sec. VI in SI for the criterion to choose the threshold).

Evaluation metrics. We use three metrics [46] to characterize the performance of our reconstruction framework: the area under the receiver operating characteristic curve (AUROC), the area under the precision-recall curve (AUPR), and the Success rate.

To define AUROC and AUPR, it is necessary to calculate three basic quantities: TPR (true positive rate), FPR (false positive rate), and Recall [46]. In particular, TPR is defined as

$$\text{TPR}(l) = \frac{\text{TP}(l)}{T}, \quad (11)$$

where l is the cut-off index in the list of the predicted links, $\text{TP}(l)$ is the number of true positives in the top l predictions in the link list, and T is the number of positives.

FPR is defined as

$$\text{FPR}(l) = \frac{\text{FP}(l)}{Q}, \quad (12)$$

where $\text{FP}(l)$ is the number of false positives in the top l entries in the predicted link list, and Q is the number of negatives by the golden standard.

Recall and Precision are defined as

$$\text{Recall}(l) = \text{TPR}(l) = \frac{\text{TP}(l)}{T}. \quad (13)$$

and

$$\text{Precision}(l) = \frac{\text{TP}(l)}{\text{TP}(l) + \text{FP}(l)} = \frac{\text{TP}(l)}{l}, \quad (14)$$

respectively. Varying the value of l from 0 to N , we plot two sequences of points: $[\text{FPR}(l), \text{TPR}(l)]$ and $[\text{Recall}(l), \text{Precision}(l)]$. The area under the two curves correspond to the values of AUROC and AUPR, respectively. For perfect reconstruction, we have AUROC=1 and AUPR=1. In the worst case (completely random), we have AUROC=0.5 and AUPR=T/2N.

Let n_1 and n_2 be the numbers of the actual and nonexistent links in the network, respectively, n_3 and n_4 be the numbers of the predicted existent and nonexistent links. The Success rates for actual links (SREL) and nonexistent links (SRNL) are defined as n_3/n_1 and n_4/n_2 , respectively. The normalized Success rate is $\sqrt{\text{SREL} \times \text{SRNL}}$ [25].

ACKNOWLEDGMENTS

This work was supported by NSFC under Grant Nos. 61473001 and 11331009. YCL is supported by ONR through Grant No. N00014-16-1-2828. XL is supported by the National Science Fund for Distinguished Young Scholars of China (No. 61425019) and NSFC under Grant No. 71731004.

Supplementary Information for
Data based reconstruction of complex multiplex networks

Chuang Ma, Han-Shuang Chen, Xiang Li, Ying-Cheng Lai, and Hai-Feng Zhang

CONTENTS

Introduction	2
Results	2
Discussion	7
Methods	7
Acknowledgments	10
I. UAU-SIS dynamics on multiplex networks with a double-layer structure	12
II. Transition tree of UAU-SIS model	12
III. Likelihood function	14
IV. Reconstruction of virtual layer	15
V. Reconstruction of physical layer	16
VI. Selection of threshold value for identification of existent links	19
VII. Reconstruction of duplex networks with heterogeneous rates of spreading dynamics	19
References	19

I. UAU-SIS DYNAMICS ON MULTIPLEX NETWORKS WITH A DOUBLE-LAYER STRUCTURE

The UAU-SIS model was originally articulated to describe the competition between social awareness and disease spreading on double layer networks, where the physical contact layer supports an epidemic process and the virtual contact layer supports awareness diffusion [40]. The two layers share exactly the same set of nodes but their connection patterns are different. Spreading of awareness in the virtual layer is described by the conventional SIS dynamics: an unaware node (U) is informed by an aware neighbor (A) with probability λ , and an aware node can lose awareness and returns to the U state with probability δ . Epidemic dynamics in the physical layer are also of the SIS type, where an infected (I) node can infect its susceptible (S) neighbors with probability β , and an I -node returns to the S state with probability μ . There is interlayer interaction, which can be described, as follows. An I -node in the physical layer is automatically aware of the infection and changes to the A state immediately in the virtual layer. If an S -node is in the A state, its ability to infect is reduced by a factor $0 \leq \gamma < 1$. Figure 6 presents a schematic illustration of the double layer network with the described interacting dynamical processes.

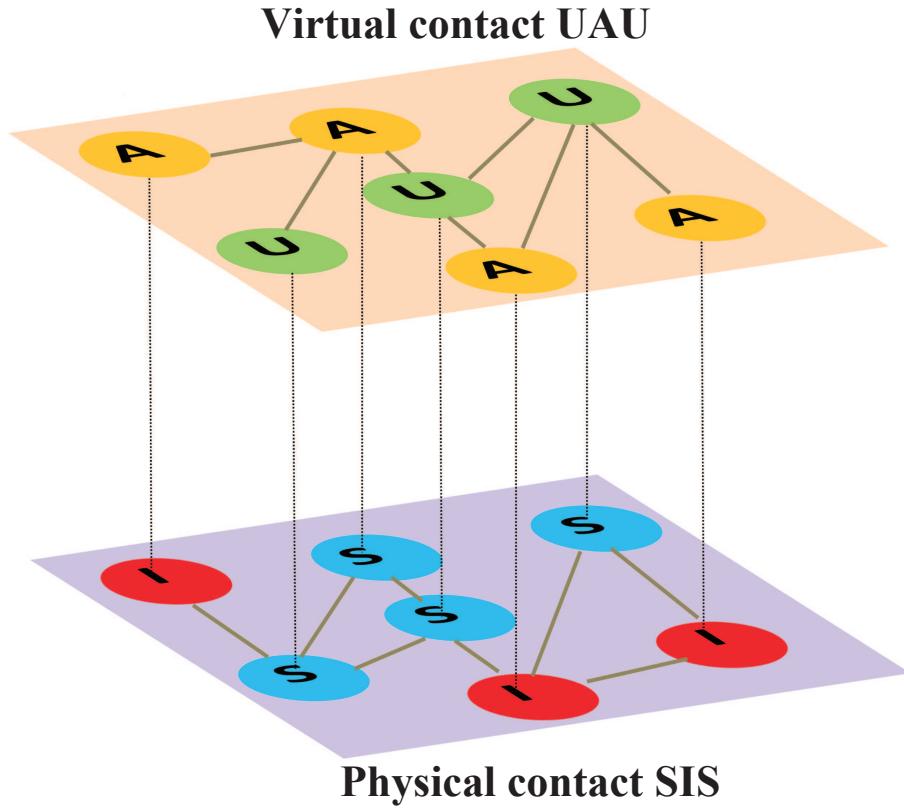


FIG. 6. **Schematic illustration a multiplex network with a double layer interacting structure.** The upper layer (virtual contact) supports awareness diffusion, where each node has two possible states: unaware (U) or aware (A). The bottom layer (physical contact) hosts epidemic spreading dynamics, where a node can be in the susceptible (S) or infected (I) state.

II. TRANSITION TREE OF UAU-SIS MODEL

Let $_{U\beta}$ and $_{A\beta} = \gamma \cdot {}_{U\beta}$ be the infection rates of unaware and aware S -node, respectively. For node i , $\sum_{j \neq i} a_j^{i-1} s_t^j$ and $\sum_{j \neq i} b_j^{i-2} s_t^j$ are the numbers of A -neighbors in the virtual layer and I -neighbors in the physical layer, respectively.

Three probabilities are needed to describe the network spreading dynamics: (1) r_i^t , the probability that node i is not informed by any neighbor, (2) ${}_{Uq_t^i}$, the probability that node i is not infected by any neighbor if i was unaware, and

(3) ${}_A q_t^i$, the probability that node i is not infected by any neighbor if i was aware. In the absence of any dynamical correlation, the three probabilities are given by

$$\begin{aligned} r_t^i &= (1 - \lambda^i)^{\sum_{j \neq i} a_j^{i,1} s_t^j}, \\ {}_U q_t^i &= (1 - {}_U \beta^i)^{\sum_{j \neq i} b_j^{i,2} s_t^j}, \\ {}_A q_t^i &= (1 - {}_A \beta^i)^{\sum_{j \neq i} b_j^{i,2} s_t^j}. \end{aligned} \quad (15)$$

A tacit assumption in Ref. [40] is that diffusion of awareness in the virtual layer occurs before epidemic spreading in the physical layer. In our work, we do not require that the two types of spreading dynamics occur in any particular order. Figure 7 presents the transition probability tree of the UAU-SIS coupling dynamics on the duplex networks.

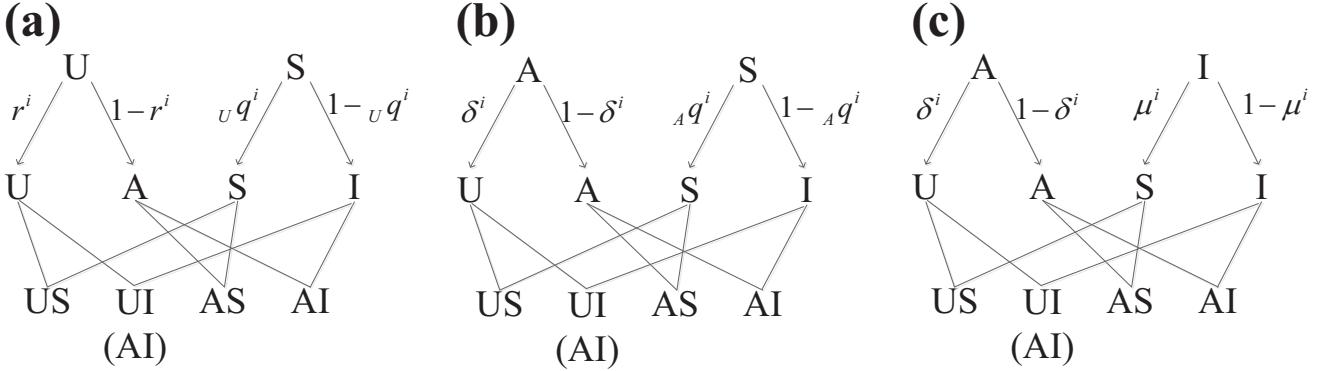


FIG. 7. **Transition probability tree of coupled UAU-SIS dynamics on duplex networks.** The notations are: AI - aware and infected, UI - unaware and infected (redundant to the AI state), AS - aware and susceptible, and US - unaware and susceptible.

From Fig. 7 and Eq. (15), we have that the probabilities of node i being in the US, AS and AI states at $t + 1$ when it is in the US state at time t are

$$\begin{aligned} P^{US \rightarrow US} &= r_t^i {}_U q_t^i, \\ P^{US \rightarrow AS} &= (1 - r_t^i) {}_U q_t^i, \\ P^{US \rightarrow AI} &= r_t^i (1 - {}_U q_t^i) + (1 - r_t^i) (1 - {}_U q_t^i) = 1 - {}_U q_t^i. \end{aligned} \quad (16)$$

The probabilities of node i being in the US, AS and AI states at $t + 1$ when it is in the AS state at time t are

$$\begin{aligned} P^{AS \rightarrow US} &= \delta^i {}_A q_t^i, \\ P^{AS \rightarrow AS} &= (1 - \delta^i) {}_A q_t^i, \\ P^{AS \rightarrow AI} &= \delta^i (1 - {}_A q_t^i) + (1 - \delta^i) (1 - {}_A q_t^i) = 1 - {}_A q_t^i. \end{aligned} \quad (17)$$

The probabilities for node i to be in the US, AS and AI states at $t + 1$ if it is in the AI state at time t are

$$\begin{aligned} P^{AI \rightarrow US} &= \delta^i \mu^i, \\ P^{AI \rightarrow AS} &= (1 - \delta^i) \mu^i, \\ P^{AI \rightarrow AI} &= 1 - \mu^i. \end{aligned} \quad (18)$$

III. LIKELIHOOD FUNCTION

The likelihood function of a node can be written in the following compact form:

$$\begin{aligned} & P \left(\left\{ {}^1 s_{t_m+1}^i, {}^2 s_{t_m+1}^i \right\}_{m=1 \dots M} \mid \left\{ {}^1 s_{t_m}^j, {}^2 s_{t_m}^j \right\}_{j=1 \dots N, m=1 \dots M}, \mathbf{a}^i, \mathbf{b}^i, \lambda^i, {}_U \beta^i, {}_A \beta^i, \delta^i, \mu^i \right) \\ & = \prod_m \left[\left(r_{t_m U}^i q_{t_m}^i \right)^{\left(1 - {}^1 s_{t_m+1}^i \right) \left(1 - {}^2 s_{t_m+1}^i \right)} \left(\left(1 - r_{t_m}^i \right) {}_U q_{t_m}^i \right)^{{}^1 s_{t_m+1}^i \left(1 - {}^2 s_{t_m+1}^i \right)} \left(1 - {}_U q_{t_m}^i \right)^{{}^1 s_{t_m+1}^i {}^2 s_{t_m+1}^i} \right]^{\left(1 - {}^1 s_{t_m}^i \right) \left(1 - {}^2 s_{t_m}^i \right)} \\ & \quad \left[\left(\delta^i {}_A q_{t_m}^i \right)^{\left(1 - {}^1 s_{t_m+1}^i \right) \left(1 - {}^2 s_{t_m+1}^i \right)} \left(\left(1 - \delta^i \right) {}_A q_{t_m}^i \right)^{{}^1 s_{t_m+1}^i \left(1 - {}^2 s_{t_m+1}^i \right)} \left(1 - {}_A q_{t_m}^i \right)^{{}^1 s_{t_m+1}^i {}^2 s_{t_m+1}^i} \right]^{{}^1 s_{t_m}^i \left(1 - {}^2 s_{t_m}^i \right)} \\ & \quad \left[\left(\delta^i \mu^i \right)^{\left(1 - {}^1 s_{t_m+1}^i \right) \left(1 - {}^2 s_{t_m+1}^i \right)} \left(\left(1 - \delta^i \right) \mu^i \right)^{{}^1 s_{t_m+1}^i \left(1 - {}^2 s_{t_m+1}^i \right)} \left(1 - \mu^i \right)^{{}^1 s_{t_m+1}^i {}^2 s_{t_m+1}^i} \right]^{{}^1 s_{t_m}^i {}^2 s_{t_m}^i}. \end{aligned} \quad (19)$$

In the UAU-SIS model, we assume that a node, once it is infected in the physical layer, enters into the A state immediately in the virtual layer: ${}^2 s_t^i = 1$ indicates ${}^1 s_t^i = 1$ at any time t . As a result, we have ${}^1 s_{t_m}^i {}^2 s_{t_m}^i = {}^2 s_{t_m}^i$ and ${}^1 s_{t_m+1}^i {}^2 s_{t_m+1}^i = {}^2 s_{t_m+1}^i$. Note that a node in the U state in the virtual layer cannot be in the I state in the physical layer, i.e., ${}^1 s_t^i = 0$ indicates ${}^2 s_t^i = 0$ at any time t , which leads to

$$\begin{aligned} \left(1 - {}^1 s_{t_m}^i \right) \left(1 - {}^2 s_{t_m}^i \right) &= 1 - {}^1 s_{t_m}^i \\ \left(1 - {}^1 s_{t_m+1}^i \right) \left(1 - {}^2 s_{t_m+1}^i \right) &= 1 - {}^1 s_{t_m+1}^i. \end{aligned}$$

Equation (19) appears complicated, but it can be reduced to some simple form. For example, assuming that node i at t_m is in the US state: ${}^1 s_{t_m}^i = 0$ and ${}^2 s_{t_m}^i = 0$, we need to retain only one term in the product:

$$\left[\left(r_{t_m U}^i q_{t_m}^i \right)^{\left(1 - {}^1 s_{t_m+1}^i \right) \left(1 - {}^2 s_{t_m+1}^i \right)} \left(\left(1 - r_{t_m}^i \right) {}_U q_{t_m}^i \right)^{{}^1 s_{t_m+1}^i \left(1 - {}^2 s_{t_m+1}^i \right)} \left(1 - {}_U q_{t_m}^i \right)^{{}^1 s_{t_m+1}^i {}^2 s_{t_m+1}^i} \right],$$

which can be further reduced to $\left(1 - r_{t_m}^i \right) {}_U q_{t_m}^i$ if ${}^1 s_{t_m+1}^i = 1$ and ${}^2 s_{t_m+1}^i = 0$ (i.e., in the AS state at the next time step). This corresponds to the transition probability from the US state to the AS state:

$$P^{US \rightarrow AS} = \left(1 - r_t^i \right) {}_U q_t^i$$

in Eq. (16). In general, Eq. (19) contains all the transition probabilities in Eqs. (16)-(18).

After some algebra, we obtain the logarithmic form of Eq. (19) as

$$L \left(\mathbf{a}^i, \mathbf{b}^i, \lambda^i, {}_U \beta^i, {}_A \beta^i, \delta^i, \mu^i \right) = L_0 \left(\delta^i, \mu^i \right) + L_1 \left(\mathbf{a}^i, \lambda^i \right) + L_2 \left(\mathbf{b}^i, {}_U \beta^i, {}_A \beta^i \right), \quad (20)$$

where

$$L_0 \left(\delta^i, \mu^i \right) = \sum_m \left[\begin{aligned} & {}^1 s_{t_m}^i \left(1 - {}^1 s_{t_m+1}^i \right) \ln \left(\delta^i \right) + {}^1 s_{t_m}^i {}^1 s_{t_m+1}^i \left(1 - {}^2 s_{t_m+1}^i \right) \ln \left(1 - \delta^i \right) \\ & + {}^2 s_{t_m}^i \left(1 - {}^2 s_{t_m+1}^i \right) \ln \left(\mu^i \right) + {}^2 s_{t_m}^i {}^2 s_{t_m+1}^i \ln \left(1 - \mu^i \right) \end{aligned} \right]. \quad (21)$$

Note that Eq. (21) does not rely on any information about the network structure. The quantity that does contain the information is $L_1 \left(\mathbf{a}^i, \lambda^i \right)$, which depends on the connectivity of node i in the virtual layer. It can be written as

$$L_1 \left(\mathbf{a}^i, \lambda^i \right) = \sum_m \left[{}^1 X_{t_m}^i \ln \left(\left(1 - \lambda^i \right)^{\sum_{j \neq i} a_j^i {}^1 s_{t_m}^j} \right) + {}^1 Y_{t_m}^i \ln \left(1 - \left(1 - \lambda^i \right)^{\sum_{j \neq i} a_j^i {}^1 s_{t_m}^j} \right) \right], \quad (22)$$

where

$$\begin{aligned} {}^1 X_{t_m}^i &= \left(1 - {}^1 s_{t_m}^i \right) \left(1 - {}^1 s_{t_m+1}^i \right) \text{ and} \\ {}^1 Y_{t_m}^i &= \left(1 - {}^1 s_{t_m}^i \right) \left(1 - {}^2 s_{t_m+1}^i \right) {}^1 s_{t_m+1}^i. \end{aligned}$$

Similarly, the quantity $L_2 \left(\mathbf{b}^i, {}_U \beta^i, {}_A \beta^i \right)$ that depends on the connectivity of node i in the physical layer is given by

$$\begin{aligned} & L_2 \left(\mathbf{b}^i, {}_U \beta^i, {}_A \beta^i \right) = \\ & \sum_m \left[{}^2 {}_U X_{t_m}^i \ln \left(\left(1 - {}_U \beta^i \right)^{\sum_{j \neq i} b_j^i {}^2 s_{t_m}^j} \right) + {}^2 {}_U Y_{t_m}^i \ln \left(1 - \left(1 - {}_U \beta^i \right)^{\sum_{j \neq i} b_j^i {}^2 s_{t_m}^j} \right) \right] \\ & + \sum_m \left[{}^2 {}_A X_{t_m}^i \ln \left(\left(1 - {}_A \beta^i \right)^{\sum_{j \neq i} b_j^i {}^2 s_{t_m}^j} \right) + {}^2 {}_A Y_{t_m}^i \ln \left(1 - \left(1 - {}_A \beta^i \right)^{\sum_{j \neq i} b_j^i {}^2 s_{t_m}^j} \right) \right] \end{aligned} \quad (23)$$

where

$$\begin{aligned} {}^2_U X_{t_m}^i &= (1 - {}^1 s_{t_m}^i) (1 - {}^2 s_{t_m+1}^i), \\ {}^2_U Y_{t_m}^i &= (1 - {}^1 s_{t_m}^i)^2 s_{t_m+1}^i, \\ {}^2_A X_{t_m}^i &= {}^1 s_{t_m}^i (1 - {}^2 s_{t_m}^i) (1 - {}^2 s_{t_m+1}^i), \text{ and} \\ {}^2_A Y_{t_m}^i &= {}^1 s_{t_m}^i (1 - {}^2 s_{t_m}^i)^2 s_{t_m+1}^i. \end{aligned}$$

In principle, Eq. (20) indicates that one can maximize L_1 and L_2 with respect to a_j^i and b_j^i , respectively, to uncover the connectivity of node i . However, the conventional maximization process leads to equations that cannot be solved because the quantity a_j^i (b_j^i) appears in the exponential term and the values of λ^i (or $U\beta^i$) are unknown. We exploit the mean-field approximation to maximize L_1 and L_2 .

IV. RECONSTRUCTION OF VIRTUAL LAYER

To infer the neighbors of node i in the virtual layer, we impose the mean-field approximation on L_1 :

$$\sum_{j \neq i} {}^1 s_{t_m}^j a_j^i \approx \frac{{}^1 k^i}{N-1} {}^1 \theta_{t_m}^i, \quad (24)$$

where N and ${}^1 k^i$ are the number of nodes and the degree of node i in the virtual layer, respectively, and the number of A-nodes in the virtual layer (excluding node i itself) is ${}^1 \theta_{t_m}^i = \sum_{j \neq i} {}^1 s_{t_m}^j$. A new unknown parameter ${}^1 k^i$ emerges in

Eq. (22) when we substitute Eq. (24) into Eq. (22). To simplify the analysis, we let

$${}^1 \gamma^i = (1 - \lambda^i)^{\frac{{}^1 k^i}{N-1}}, \quad (25)$$

to obtain

$$(1 - \lambda^i)^{\sum_{j \neq i} a_j^i} {}^1 s_{t_m}^j = (1 - \lambda^i)^{\frac{{}^1 k^i}{N-1}} {}^1 \theta_{t_m}^i = ({}^1 \gamma^i)^{{}^1 \theta_{t_m}^i}.$$

Equation (22) can then be written concisely as

$$\hat{L}_1({}^1 \gamma^i) = \sum_m \left[{}^1 X_{t_m}^i \ln \left(({}^1 \gamma^i)^{{}^1 \theta_{t_m}^i} \right) + {}^1 Y_{t_m}^i \ln \left(1 - ({}^1 \gamma^i)^{{}^1 \theta_{t_m}^i} \right) \right]. \quad (26)$$

Differentiating $\hat{L}_1({}^1 \gamma^i)$ with respect to ${}^1 \gamma^i$ and setting it to zero, we get

$$\sum_m {}^1 Y_{t_m}^i {}^1 \theta_{t_m}^i \frac{({}^1 \gamma^i)^{{}^1 \theta_{t_m}^i}}{1 - ({}^1 \gamma^i)^{{}^1 \theta_{t_m}^i}} = \sum_m {}^1 X_{t_m}^i {}^1 \theta_{t_m}^i,$$

leading to ${}^1 \gamma^i = {}^1 \tilde{\gamma}^i$.

Treating a_l^i as a continuous variance, we can further differentiate Eq. (22) with respect to a_l^i and set it to zero, which gives

$$\sum_m {}^1 Y_{t_m}^i {}^1 s_{t_m}^l \frac{(1 - \lambda^i)^{\sum_{j \neq i} a_j^i} {}^1 s_{t_m}^j}{1 - (1 - \lambda^i)^{\sum_{j \neq i} a_j^i} {}^1 s_{t_m}^j} = \sum_m {}^1 X_{t_m}^i {}^1 s_{t_m}^l. \quad (27)$$

To obtain analytical solutions of Eq. (27) is not feasible due to its nonlinear and high-dimensional nature. We thus resort to the first-order Taylor expansion. In particular, we expand $a^x/(1-a^x)$ in the limit $x \rightarrow x_0$ to obtain

$$\frac{a^x}{1-a^x} \approx \frac{a^{x_0}}{1-a^{x_0}} + \frac{a^{x_0} \ln a}{(1-a^{x_0})^2} (x - x_0) = \frac{a^{x_0}}{1-a^{x_0}} - \frac{a^{x_0} \ln a^{x_0}}{(1-a^{x_0})^2} + \frac{a^{x_0} \ln a}{(1-a^{x_0})^2} x. \quad (28)$$

Setting $x = \sum_{j \neq i} a_j^i {}^1 s_{t_m}^j$, $a = 1 - \lambda^i$, and $x_0 = \frac{{}^1 k^i}{N-1} {}^1 \theta_{t_m}^i$ [Eq. (24) implies that $x \approx x_0$], we have

$$a^{x_0} = ({}^1 \tilde{\gamma}^i)^{{}^1 \theta_{t_m}^i}$$

from Eq. (25). That is, the term

$$\frac{(1 - \lambda^i)^{\sum_{j \neq i} a_j^{i-1} s_{t_m}^j}}{1 - (1 - \lambda^i)^{\sum_{j \neq i} a_j^{i-1} s_{t_m}^j}}$$

in Eq. (27) can be expressed as a linear form based on Eq. (24.)

Collecting all the approximations, we transform Eq. (27) into a solvable linear system as

$$\sum_m {}^1 Y_{t_m}^i {}^1 G_{t_m}^i {}^1 s_{t_m}^l \ln(1 - \lambda^i) \sum_{j \neq i} a_j^{i-1} s_{t_m}^j = \sum_m ({}^1 X_{t_m}^i - {}^1 Y_{t_m}^i {}^1 F_{t_m}^i) {}^1 s_{t_m}^l, \quad (29)$$

where

$$\begin{aligned} {}^1 F_{t_m}^i &= \frac{({}^1 \tilde{\gamma}^i)^{1 \theta_{t_m}^i}}{1 - ({}^1 \tilde{\gamma}^i)^{1 \theta_{t_m}^i}} - \frac{({}^1 \tilde{\gamma}^i)^{1 \theta_{t_m}^i}}{\left(1 - ({}^1 \tilde{\gamma}^i)^{1 \theta_{t_m}^i}\right)^2} {}^1 \theta_{t_m}^i \ln {}^1 \tilde{\gamma}^i \quad \text{and} \\ {}^1 G_{t_m}^i &= \frac{({}^1 \tilde{\gamma}^i)^{1 \theta_{t_m}^i}}{\left(1 - ({}^1 \tilde{\gamma}^i)^{1 \theta_{t_m}^i}\right)^2}. \end{aligned}$$

Setting ${}^1 \Phi_{t_m}^i = {}^1 Y_{t_m}^i {}^1 G_{t_m}^i$ and ${}^1 \Gamma_{t_m}^i = {}^1 X_{t_m}^i - {}^1 Y_{t_m}^i {}^1 F_{t_m}^i$, we express Eq. (29) as

$$\begin{bmatrix} \sum_m {}^1 \Phi_{t_m}^i {}^1 s_{t_m}^1 {}^1 s_{t_m}^1 & \dots & \sum_m {}^1 \Phi_{t_m}^i {}^1 s_{t_m}^1 {}^1 s_{t_m}^{i-1} & \sum_m {}^1 \Phi_{t_m}^i {}^1 s_{t_m}^1 {}^1 s_{t_m}^{i+1} & \dots & \sum_m {}^1 \Phi_{t_m}^i {}^1 s_{t_m}^1 {}^1 s_{t_m}^N \\ \vdots & & \vdots & \vdots & & \vdots \\ \sum_m {}^1 \Phi_{t_m}^i {}^1 s_{t_m}^{i-1} {}^1 s_{t_m}^1 & \dots & \sum_m {}^1 \Phi_{t_m}^i {}^1 s_{t_m}^{i-1} {}^1 s_{t_m}^{i-1} & \sum_m {}^1 \Phi_{t_m}^i {}^1 s_{t_m}^{i-1} {}^1 s_{t_m}^{i+1} & \dots & \sum_m {}^1 \Phi_{t_m}^i {}^1 s_{t_m}^{i-1} {}^1 s_{t_m}^N \\ \sum_m {}^1 \Phi_{t_m}^i {}^1 s_{t_m}^{i+1} {}^1 s_{t_m}^1 & \dots & \sum_m {}^1 \Phi_{t_m}^i {}^1 s_{t_m}^{i+1} {}^1 s_{t_m}^{i-1} & \sum_m {}^1 \Phi_{t_m}^i {}^1 s_{t_m}^{i+1} {}^1 s_{t_m}^{i+1} & \dots & \sum_m {}^1 \Phi_{t_m}^i {}^1 s_{t_m}^{i+1} {}^1 s_{t_m}^N \\ \vdots & & \vdots & \vdots & & \vdots \\ \sum_m {}^1 \Phi_{t_m}^i {}^1 s_{t_m}^N {}^1 s_{t_m}^1 & \dots & \sum_m {}^1 \Phi_{t_m}^i {}^1 s_{t_m}^N {}^1 s_{t_m}^{i-1} & \sum_m {}^1 \Phi_{t_m}^i {}^1 s_{t_m}^N {}^1 s_{t_m}^{i+1} & \dots & \sum_m {}^1 \Phi_{t_m}^i {}^1 s_{t_m}^N {}^1 s_{t_m}^N \end{bmatrix} \times \begin{bmatrix} a_1^i \ln(1 - \lambda^i) \\ \vdots \\ a_{i-1}^i \ln(1 - \lambda^i) \\ a_{i+1}^i \ln(1 - \lambda^i) \\ \vdots \\ a_N^i \ln(1 - \lambda^i) \end{bmatrix} = \begin{bmatrix} \sum_m {}^1 \Gamma_{t_m}^i {}^1 s_{t_m}^1 \\ \vdots \\ \sum_m {}^1 \Gamma_{t_m}^i {}^1 s_{t_m}^{i-1} \\ \sum_m {}^1 \Gamma_{t_m}^i {}^1 s_{t_m}^{i+1} \\ \vdots \\ \sum_m {}^1 \Gamma_{t_m}^i {}^1 s_{t_m}^N \end{bmatrix}. \quad (30)$$

The matrix on the left side (labeled as Λ) and the vector (labeled as ζ) on the right side of Eq. (30) can be calculated from the time series of the nodal states. The vector

$$\boldsymbol{\eta} = [a_1^i \ln(1 - \lambda^i), \dots, a_{i-1}^i \ln(1 - \lambda^i), a_{i+1}^i \ln(1 - \lambda^i), \dots, a_N^i \ln(1 - \lambda^i)]^T$$

can then be solved, where T denotes transpose. Note that the quantity $\ln(1 - \lambda^i) < 0$ is a constant even though λ^i is not given, implying that the value of $-a_j^i \ln(1 - \lambda^i)$ is positively large for $a_j^i = 1$ and near zero for $a_j^i = 0$. As a result, the neighbors of node i in the virtual layer can be ascertained through the solution of the vector $\boldsymbol{\eta}$.

V. RECONSTRUCTION OF PHYSICAL LAYER

The mean-field approximation for the physical layer is

$$\sum_{j \neq i} {}^2 s_{t_m}^j b_j^i \approx \frac{2k^i}{N-1} {}^2 \theta_{t_m}^i, \quad (31)$$

where ${}^2k^i$ is the degree of node i and ${}^2\theta_{t_m}^i = \sum_{j \neq i} {}^2s_{t_m}^j$ is the number of I-nodes in the physical layer (excluding node i itself). We set

$$\begin{aligned} {}_U^2\gamma^i &= (1 - {}_U\beta^i)^{\frac{{}^2k^i}{N-1}} \text{ and} \\ {}_A^2\gamma^i &= (1 - {}_A\beta^i)^{\frac{{}^2k^i}{N-1}} \end{aligned} \quad (32)$$

and write Eq. (23) concisely as

$$\hat{L}_2({}_U^2\gamma^i, {}_A^2\gamma^i) = \sum_m \left\{ \begin{array}{l} \left[{}_U^2X_{t_m}^i \ln \left(({}_U^2\gamma^i)^{{}^2\theta_{t_m}^i} \right) + {}_U^2Y_{t_m}^i \ln \left(1 - ({}_U^2\gamma^i)^{{}^2\theta_{t_m}^i} \right) \right] \\ + \left[{}_A^2X_{t_m}^i \ln \left(({}_A^2\gamma^i)^{{}^2\theta_{t_m}^i} \right) + {}_A^2Y_{t_m}^i \ln \left(1 - ({}_A^2\gamma^i)^{{}^2\theta_{t_m}^i} \right) \right] \end{array} \right\}. \quad (33)$$

Taking the derivatives of \hat{L}_2 with respect to ${}_U^2\gamma^i$ and ${}_A^2\gamma^i$ and setting them to zero, we get

$$\sum_m {}_U^2Y_{t_m}^i {}^2\theta_{t_m}^i \frac{({}_U^2\gamma^i)^{{}^2\theta_{t_m}^i}}{1 - ({}_U^2\gamma^i)^{{}^2\theta_{t_m}^i}} = \sum_m {}_U^2X_{t_m}^i {}^2\theta_{t_m}^i \quad (34)$$

and

$$\sum_m {}_A^2Y_{t_m}^i {}^2\theta_{t_m}^i \frac{({}_A^2\gamma^i)^{{}^2\theta_{t_m}^i}}{1 - ({}_A^2\gamma^i)^{{}^2\theta_{t_m}^i}} = \sum_m {}_A^2X_{t_m}^i {}^2\theta_{t_m}^i, \quad (35)$$

which give

$$\begin{aligned} {}_U^2\gamma^i &= {}_U^2\tilde{\gamma}^i \text{ and} \\ {}_A^2\gamma^i &= {}_A^2\tilde{\gamma}^i. \end{aligned}$$

Similar to the mean-field analysis of the virtual layer, we differentiate Eq. (23) with respect to b_l^i and set it to zero:

$$\begin{aligned} \frac{\partial L_2}{\partial b_l^i} &= \sum_m \left[\ln(1 - {}_U\beta^i) {}_U^2X_{t_m}^i {}^2s_{t_m}^l - \ln(1 - {}_U\beta^i) {}_U^2Y_{t_m}^i {}^2s_{t_m}^l \frac{(1 - {}_U\beta^i)^{\sum_j b_j^i {}^2s_{t_m}^j}}{1 - (1 - {}_U\beta^i)^{\sum_j b_j^i {}^2s_{t_m}^j}} \right] \\ &\quad + \sum_m \left[\ln(1 - {}_A\beta^i) {}_A^2X_{t_m}^i {}^2s_{t_m}^l - \ln(1 - {}_A\beta^i) {}_A^2Y_{t_m}^i {}^2s_{t_m}^l \frac{(1 - {}_A\beta^i)^{\sum_j b_j^i {}^2s_{t_m}^j}}{1 - (1 - {}_A\beta^i)^{\sum_j b_j^i {}^2s_{t_m}^j}} \right] = 0, \end{aligned}$$

which gives

$$\begin{aligned} \sum_m &\left\{ \ln(1 - {}_U\beta^i) {}^2s_{t_m}^l {}_U^2Y_{t_m}^i \frac{(1 - {}_U\beta^i)^{\sum_j b_j^i {}^2s_{t_m}^j}}{1 - (1 - {}_U\beta^i)^{\sum_j b_j^i {}^2s_{t_m}^j}} + \ln(1 - {}_A\beta^i) {}^2s_{t_m}^l {}_A^2Y_{t_m}^i \frac{(1 - {}_A\beta^i)^{\sum_j b_j^i {}^2s_{t_m}^j}}{1 - (1 - {}_A\beta^i)^{\sum_j b_j^i {}^2s_{t_m}^j}} \right\} \\ &= \sum_m \{ \ln(1 - {}_U\beta^i) {}_U^2X_{t_m}^i {}^2s_{t_m}^l + \ln(1 - {}_A\beta^i) {}_A^2X_{t_m}^i {}^2s_{t_m}^l \}. \end{aligned} \quad (36)$$

Setting

$$\rho = \frac{\ln {}_U^2\tilde{\gamma}^i}{\ln {}_A^2\tilde{\gamma}^i} = \frac{\ln(1 - {}_U\beta^i)}{\ln(1 - {}_A\beta^i)},$$

we can further simplify Eq. (36) as

$$\begin{aligned} \sum_m &\left[\rho {}^2s_{t_m}^l {}_U^2Y_{t_m}^i \frac{(1 - {}_U\beta^i)^{\sum_j b_j^i {}^2s_{t_m}^j}}{1 - (1 - {}_U\beta^i)^{\sum_j b_j^i {}^2s_{t_m}^j}} + {}^2s_{t_m}^l {}_A^2Y_{t_m}^i \frac{(1 - {}_A\beta^i)^{\sum_j b_j^i {}^2s_{t_m}^j}}{1 - (1 - {}_A\beta^i)^{\sum_j b_j^i {}^2s_{t_m}^j}} \right] \\ &= \sum_m [\rho {}_U^2X_{t_m}^i {}^2s_{t_m}^l + {}_A^2X_{t_m}^i {}^2s_{t_m}^l] \end{aligned} \quad (37)$$

Using Eq. (28) and setting $x = \sum_{j \neq i} b_j^i {}^2 s_{t_m}^j$ and $x_0 = \frac{^2 k^i}{N-1} {}^2 \theta_{t_m}^i$, we obtain the following Taylor expansion:

$$\frac{(1 - {}_U \beta^i) \sum_{j \neq i} b_j^i {}^2 s_{t_m}^j}{1 - (1 - {}_U \beta^i) \sum_{j \neq i} b_j^i {}^2 s_{t_m}^j} = {}_U^2 F_{t_m}^i + {}_U^2 G_{t_m}^i \ln(1 - {}_U \beta^i) \sum_{j \neq i} b_j^i {}^2 s_{t_m}^j, \quad (38)$$

where

$$\begin{aligned} {}_U^2 F_{t_m}^i &= \frac{({}_U^2 \tilde{\gamma}^i)^{2 \theta_{t_m}^i}}{1 - ({}_U^2 \tilde{\gamma}^i)^{2 \theta_{t_m}^i}} - \frac{({}_U^2 \tilde{\gamma}^i)^{2 \theta_{t_m}^i}}{\left(1 - ({}_U^2 \tilde{\gamma}^i)^{2 \theta_{t_m}^i}\right)^2} {}^2 \theta_{t_m}^i \ln {}_U^2 \tilde{\gamma}^i \text{ and} \\ {}_U^2 G_{t_m}^i &= \frac{({}_U^2 \tilde{\gamma}^i)^{2 \theta_{t_m}^i}}{\left(1 - ({}_U^2 \tilde{\gamma}^i)^{2 \theta_{t_m}^i}\right)^2}. \end{aligned}$$

We also have

$$\frac{(1 - {}_A \beta^i) \sum_{j \neq i} b_j^i {}^2 s_{t_m}^j}{1 - (1 - {}_A \beta^i) \sum_{j \neq i} b_j^i {}^2 s_{t_m}^j} = {}_A^2 F_{t_m}^i + {}_A^2 G_{t_m}^i \ln(1 - {}_A \beta^i) \sum_{j \neq i} b_j^i {}^2 s_{t_m}^j, \quad (39)$$

where

$$\begin{aligned} {}_A^2 F_{t_m}^i &= \frac{({}_A^2 \tilde{\gamma}^i)^{2 \theta_{t_m}^i}}{1 - ({}_A^2 \tilde{\gamma}^i)^{2 \theta_{t_m}^i}} - \frac{({}_A^2 \tilde{\gamma}^i)^{2 \theta_{t_m}^i}}{\left(1 - ({}_A^2 \tilde{\gamma}^i)^{2 \theta_{t_m}^i}\right)^2} {}^2 \theta_{t_m}^i \ln {}_A^2 \tilde{\gamma}^i \text{ and} \\ {}_A^2 G_{t_m}^i &= \frac{({}_A^2 \tilde{\gamma}^i)^{2 \theta_{t_m}^i}}{\left(1 - ({}_A^2 \tilde{\gamma}^i)^{2 \theta_{t_m}^i}\right)^2}. \end{aligned}$$

With these approximations, we can transform Eq. (37) into the following linear system:

$$\begin{aligned} &\sum_m \left[(\rho {}_U^2 Y_{t_m}^i {}_U^2 G_{t_m}^i + {}_A^2 Y_{t_m}^i {}_A^2 G_{t_m}^i) {}^2 s_{t_m}^l \ln(1 - {}_A \beta^i) \sum_{j \neq i} b_j^i {}^2 s_{t_m}^j \right] \\ &= \sum_m (\rho {}_U^2 X_{t_m}^i + {}_A^2 X_{t_m}^i - \rho {}_U^2 Y_{t_m}^i {}_U^2 F_{t_m}^i - {}_A^2 Y_{t_m}^i {}_A^2 F_{t_m}^i) {}^2 s_{t_m}^l, \end{aligned} \quad (40)$$

where

$$\begin{aligned} {}^2 \Phi_{t_m}^i &= \rho {}_U^2 Y_{t_m}^i {}_U^2 G_{t_m}^i + {}_A^2 Y_{t_m}^i {}_A^2 G_{t_m}^i \text{ and} \\ {}^2 \Gamma_{t_m}^i &= \rho {}_U^2 X_{t_m}^i + {}_A^2 X_{t_m}^i - \rho {}_U^2 Y_{t_m}^i {}_U^2 F_{t_m}^i - {}_A^2 Y_{t_m}^i {}_A^2 F_{t_m}^i. \end{aligned}$$

We rewrite Eq. (40) in the following matrix form:

$$\begin{bmatrix} \sum_m^2 \Phi_{t_m}^i 2 s_{t_m}^1 2 s_{t_m}^1 & \dots & \sum_m^2 \Phi_{t_m}^i 2 s_{t_m}^1 2 s_{t_m}^{i-1} & \sum_m^2 \Phi_{t_m}^i 2 s_{t_m}^1 2 s_{t_m}^{i+1} & \dots & \sum_m^2 \Phi_{t_m}^i 2 s_{t_m}^1 2 s_{t_m}^N \\ \vdots & & \vdots & & \vdots & \vdots \\ \sum_m^2 \Phi_{t_m}^i 2 s_{t_m}^{i-1} 2 s_{t_m}^1 & \dots & \sum_m^2 \Phi_{t_m}^i 2 s_{t_m}^{i-1} 2 s_{t_m}^{i-1} & \sum_m^2 \Phi_{t_m}^i 2 s_{t_m}^{i-1} 2 s_{t_m}^{i+1} & \dots & \sum_m^2 \Phi_{t_m}^i 2 s_{t_m}^{i-1} 2 s_{t_m}^N \\ \sum_m^2 \Phi_{t_m}^i 2 s_{t_m}^{i+1} 2 s_{t_m}^1 & \dots & \sum_m^2 \Phi_{t_m}^i 2 s_{t_m}^{i+1} 2 s_{t_m}^{i-1} & \sum_m^2 \Phi_{t_m}^i 2 s_{t_m}^{i+1} 2 s_{t_m}^{i+1} & \dots & \sum_m^2 \Phi_{t_m}^i 2 s_{t_m}^{i+1} 2 s_{t_m}^N \\ \vdots & & \vdots & & \vdots & \vdots \\ \sum_m^2 \Phi_{t_m}^i 2 s_{t_m}^N 2 s_{t_m}^1 & \dots & \sum_m^2 \Phi_{t_m}^i 2 s_{t_m}^N 2 s_{t_m}^{i-1} & \sum_m^2 \Phi_{t_m}^i 2 s_{t_m}^N 2 s_{t_m}^{i+1} & \dots & \sum_m^2 \Phi_{t_m}^i 2 s_{t_m}^N 2 s_{t_m}^N \end{bmatrix} \\ \times \begin{bmatrix} b_1^i \ln(1 - {}_A\beta^i) \\ \vdots \\ b_{i-1}^i \ln(1 - {}_A\beta^i) \\ b_{i+1}^i \ln(1 - {}_A\beta^i) \\ \vdots \\ b_N^i \ln(1 - {}_A\beta^i) \end{bmatrix} = \begin{bmatrix} \sum_m^2 \Gamma_{t_m}^i 2 s_{t_m}^1 \\ \vdots \\ \sum_m^2 \Gamma_{t_m}^i 2 s_{t_m}^{i-1} \\ \sum_m^2 \Gamma_{t_m}^i 2 s_{t_m}^{i+1} \\ \vdots \\ \sum_m^2 \Gamma_{t_m}^i 2 s_{t_m}^N \end{bmatrix}. \quad (41)$$

The matrix on the left side and the vector on the right side can be obtained from time series. Solution of Eq. (41) gives the vector

$$\boldsymbol{\eta} = [b_1^i \ln(1 - {}_A\beta^i), \dots, b_{i-1}^i \ln(1 - {}_A\beta^i), b_{i+1}^i \ln(1 - {}_A\beta^i), \dots, b_N^i \ln(1 - {}_A\beta^i)]^T,$$

revealing the neighbors of node i in the physical layer.

VI. SELECTION OF THRESHOLD VALUE FOR IDENTIFICATION OF EXISTENT LINKS

For each node i , the values of $a_i^i \ln(1 - \lambda^i)$ (or of $b_i^i \ln(1 - {}_A\beta^i)$) can be obtained from Eq. (30) [or Eq. (41)]. From Figs. 8(a,b), we have that the values of $-a_l^i \ln(1 - \lambda^i)$ [or $-b_l^i \ln(1 - {}_A\beta^i)$] are unequivocally above zero for the actual links, while their values are close to zero for nonexistent links, with a gap between the two sets of values. Representing the values listed in each column as a histogram, we have that the peak centered about zero corresponds to nonexistent links and the other corresponds to existent links. A threshold value can be placed between the two peaks [25], as shown in Fig. 8(c). A pair of nodes i and l are connected if the corresponding value of $-a_l^i \ln(1 - \lambda^i)$ [$-b_l^i \ln(1 - {}_A\beta^i)$] is larger than the threshold. Take node 46 as an example. We wish to infer its neighbors in the physical layer [highlighted by the red dashed frame in Fig. 8(b)]. Figure 8(d) shows that the values larger than the threshold correspond to the existent links.

VII. RECONSTRUCTION OF DUPLEX NETWORKS WITH HETEROGENEOUS RATES OF SPREADING DYNAMICS

Figure 9 demonstrates that our framework can reconstruct duplex networks with heterogeneous rates of spreading dynamics. In particular, transmission rates λ^i and ${}_U\beta^i$ are randomly chosen from the ranges $(0.2, 0.4)$ and $(0.3, 0.5)$, respectively. The recovery rates δ^i and μ^i are randomly picked up from the ranges $(0.6, 1)$ and $(0.6, 1)$, respectively. Note that ${}_A\beta^i = 0.5 {}_U\beta^i$.

-
- [1] Friston, K. J. Bayesian estimation of dynamical systems: an application to fMRI. *NeuroImage* **16**, 513–530 (2002).
 - [2] Gardner, T. S., di Bernardo, D., Lorenz, D. & Collins, J. J. Inferring genetic networks and identifying compound mode of action via expression profiling. *Science* **301**, 102–105 (2003).
 - [3] Pipa, G. & Grün, S. Non-parametric significance estimation of joint-spike events by shuffling and resampling. *Neurocomp.* **52**, 31–37 (2003).

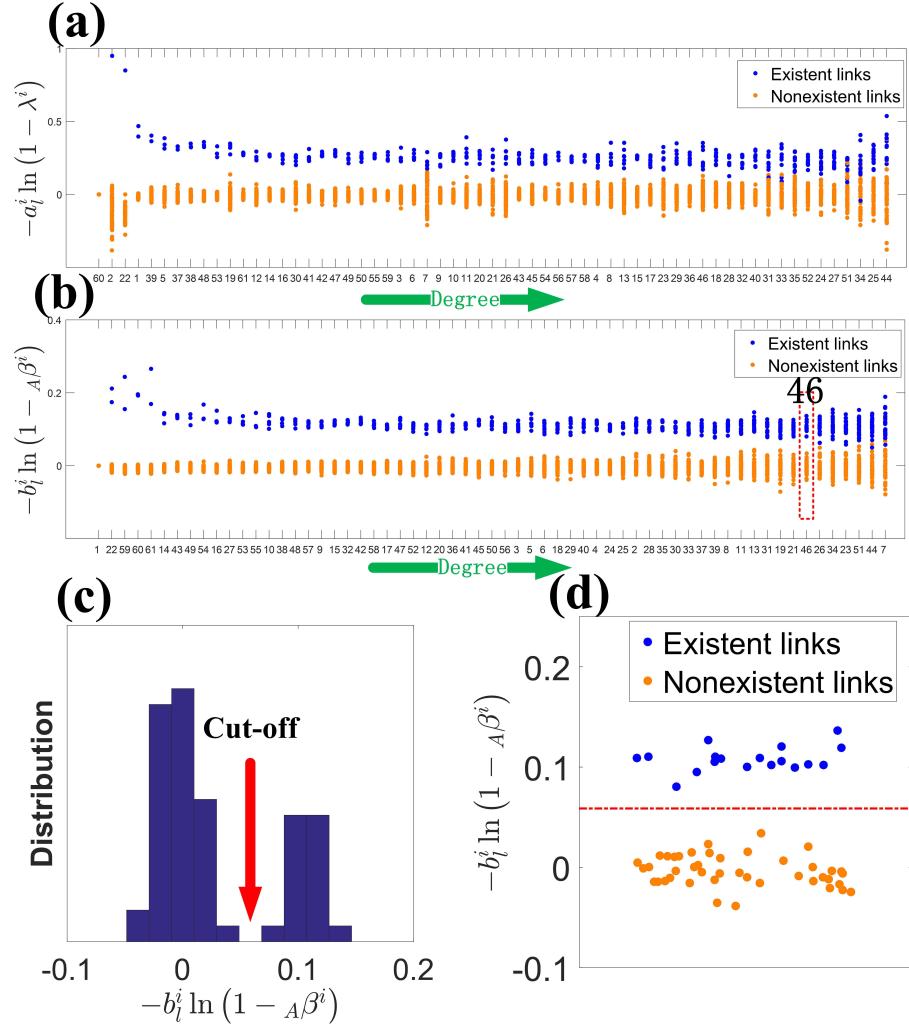


FIG. 8. Reconstruction of CS-AARHUS duplex network. (a,b) Values of $-a_l^i \ln(1 - \lambda^i)$, $i \neq l$ and $-b_l^i \ln(1 - \lambda^i)$ $i \neq l$ (b) for each node, respectively. Each column gives the connectivity of a node. The blue and orange points denote the existent and nonexistent links, respectively. (c) Illustration of the choice of the threshold with node 46 [highlighted by the red dashed frame in (b)]. Shown is the distribution of the values of $-b_l^{46} \ln(1 - \lambda^i)$ for $l \neq 46$. The peak centered about zero corresponds to nonexistent links while the other peak corresponds to existent links. A threshold can be set within the gap between the two peaks. (d) The threshold is illustrated to distinguish the actual from the nonexistent links. The length of time series is $M=30000$. Other parameters are $\lambda = 0.2$, $U\beta = 0.2$, $A\beta = 0.5U\beta$, and $\mu = \delta = 0.8$.

- [4] Brovelli, A. *et al.* Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by Granger causality. *Proc. Natl. Acad. Sci. (USA)* **101**, 9849–9854 (2004).
- [5] Eguiluz, V. M., Chialvo, D. R., Cecchi, G. A., Baliki, M. & Apkarian, A. V. Scale-free brain functional networks. *Phys. Rev. Lett.* **94**, 018102 (2005).
- [6] Yu, D., Righero, M. & Kocarev, L. Estimating topology of networks. *Phys. Rev. Lett.* **97**, 188701 (2006).
- [7] Borgard, J. & Lipson, H. Automated reverse engineering of nonlinear dynamical systems. *Proc. Natl. Acad. Sci. (USA)* **104**, 9943–9948 (2007).
- [8] Timme, M. Revealing network connectivity from response dynamics. *Phys. Rev. Lett.* **98**, 224101 (2007).
- [9] Napoletani, D. & Sauer, T. D. Reconstructing the topology of sparsely connected dynamical networks. *Phys. Rev. E* **77**, 026103 (2008).
- [10] Donges, J., Zou, Y., Marwan, N. & Kurths, J. The backbone of the climate network. *EPL (Europhys. Lett.)* **87**, 48007 (2009).
- [11] Pajevic, S. & Plenz, D. Efficient network reconstruction from dynamical cascades identifies small-world topology of neuronal avalanches. *PLoS Comp. Biol.* **5**, e1000271 (2009).
- [12] Ren, J., Wang, W.-X., Li, B. & Lai, Y.-C. Noise bridges dynamical correlation and topology in coupled oscillator networks. *Phys. Rev. Lett.* **104**, 058701 (2010).

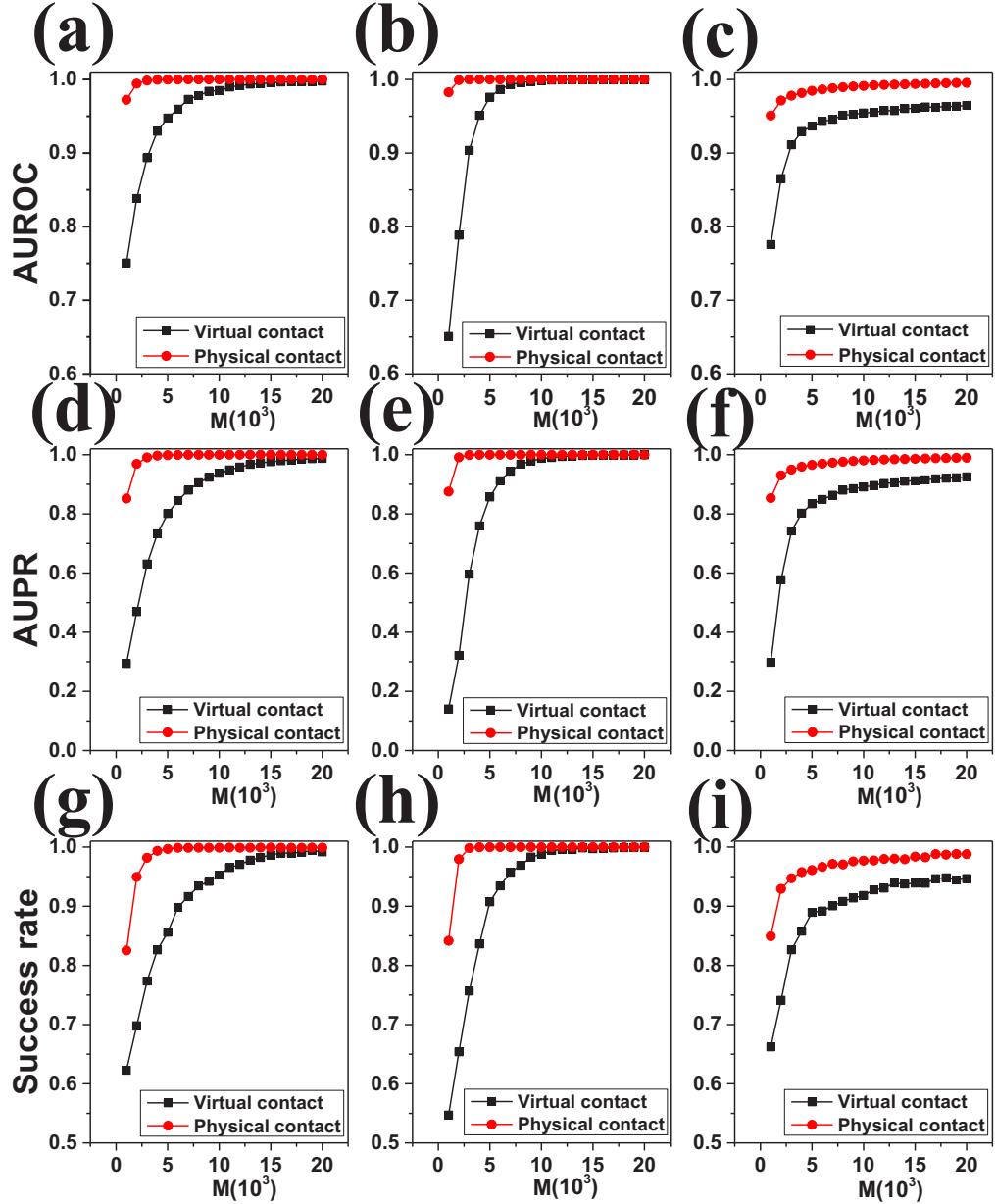


FIG. 9. Applicability of reconstruction framework to spreading dynamics with heterogeneous rates. Reconstruction accuracy versus the length M of the time series for ER-ER (left column), SW-SW (central column) and BA-BA (right column) duplex networks with heterogeneous transmission and recovery rates. The network parameters are $N = 100$ and $\langle k_1 \rangle = 4$, $\langle k_2 \rangle = 6$.

- [13] Levnajić, Z. & Pikovsky, A. Network reconstruction from random phase resetting. *Phys. Rev. Lett.* **107**, 034101 (2011).
- [14] Hempel, S., Koseska, A., Kurths, J. & Nikoloski, Z. Inner composition alignment for inferring directed networks from short time series. *Phys. Rev. Lett.* **107**, 054101 (2011).
- [15] Shandilya, S. G. & Timme, M. Inferring network topology from complex dynamics. *New J. Phys.* **13**, 013004 (2011).
- [16] Wang, W.-X., Lai, Y.-C., Grebogi, C. & Ye, J.-P. Network reconstruction based on evolutionary-game data via compressive sensing. *Phys. Rev. X* **1**, 021021 (2011).
- [17] Berry, T., Hamilton, F., Peixoto, N. & Sauer, T. Detecting connectivity changes in neuronal networks. *J. Neurosci. Meth.* **209**, 388–397 (2012).
- [18] Stetter, O., Battaglia, D., Soriano, J. & Geisel, T. Model-free reconstruction of excitatory neuronal connectivity from calcium imaging signals. *PLoS Comp. Biol.* **8**, e1002653 (2012).

- [19] Su, R.-Q., Wang, W.-X. & Lai, Y.-C. Detecting hidden nodes in complex networks from time series. *Phys. Rev. E* **85**, 065201 (2012).
- [20] Hamilton, F., Berry, T., Peixoto, N. & Sauer, T. Real-time tracking of neuronal network structure using data assimilation. *Phys. Rev. E* **88**, 052715 (2013).
- [21] Zhou, D., Xiao, Y., Zhang, Y., Xu, Z. & Cai, D. Causal and structural connectivity of pulse-coupled nonlinear networks. *Phys. Rev. Lett.* **111**, 054102 (2013).
- [22] Ching, E. S. C., Lai, P.-Y. & Leung, C. Y. Extracting connectivity from dynamics of networks with uniform bidirectional coupling. *Phys. Rev. E* **88**, 042817 (2013).
- [23] Timme, M. & Casadiego, J. Revealing networks from dynamics: an introduction. *J. Phys. A. Math. Theo.* **47**, 343001 (2014).
- [24] Su, R.-Q., Lai, Y.-C. & Wang, X. Identifying chaotic FitzHugh-Nagumo neurons using compressive sensing. *Entropy* **16**, 3889–3902 (2014).
- [25] Shen, Z.-S., Wang, W.-X., Fan, Y., Di, Z. & Lai, Y.-C. Reconstructing propagation networks with natural diversity and identifying hidden sources. *Nat. Commun.* **5**, 4323 (2014).
- [26] Ching, E. S. C., Lai, P.-Y. & Leung, C. Y. Reconstructing weighted networks from dynamics. *Phys. Rev. E* **91**, 030801 (2015).
- [27] Wang, W.-X., Lai, Y.-C. & Grebogi, C. Data based identification and prediction of nonlinear and complex dynamical systems. *Phys. Rep.* **644**, 1–76 (2016).
- [28] Casadiego, J., Nitzan, M., Hallerberg, S. & Timme, M. Model-free inference of direct network interactions from nonlinear collective dynamics. *Nat. Commun.* **8**, 2192 (2017).
- [29] Li, X. & Li, X. Reconstruction of stochastic temporal networks through diffusive arrival times. *Nat. Commun.* **8**, 15729 (2017).
- [30] Chen, Y.-Z. & Lai, Y.-C. Sparse dynamical Boltzmann machine for reconstructing complex networks with binary dynamics. *Phys. Rev. E* **97**, 032317 (2018).
- [31] Mei, G. *et al.* Compressive-sensing-based structure identification for multilayer networks. *IEEE Trans. Cybernet* **48**, 754–764 (2018).
- [32] Nitzan, M., Casadiego, J. & Timme, M. Revealing physical interaction networks from statistics of collective dynamics. *Sci. Adv.* **3** (2017).
- [33] Ma, C., Zhang, H.-F. & Lai, Y.-C. Reconstructing complex networks without time series. *Phys. Rev. E* **96**, 022320 (2017).
- [34] Wang, W.-X., Yang, R., Lai, Y.-C., Kovanis, V. & Harrison, M. A. F. Time-series-based prediction of complex oscillator networks via compressive sensing. *EPL (Europhys. Lett.)* **94**, 48006 (2011).
- [35] Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E. & Havlin, S. Catastrophic cascade of failures in interdependent networks. *Nature* **464**, 1025 (2010).
- [36] Gao, J., Buldyrev, S. V., Stanley, H. E. & Havlin, S. Networks formed from interdependent networks. *Nat. Phys.* **8**, 40 (2012).
- [37] De Domenico, M. *et al.* Mathematical formulation of multilayer networks. *Phys. Rev. X* **3**, 041022 (2013).
- [38] Kivelä, M. *et al.* Multilayer networks. *J. Complex Net.* **2**, 203–271 (2014).
- [39] De Domenico, M., Granell, C., Porter, M. A. & Arenas, A. The physics of spreading processes in multilayer networks. *Nat. Phys.* **12**, 901 (2016).
- [40] Granell, C., Gómez, S. & Arenas, A. Dynamical interplay between awareness and epidemic spreading in multiplex networks. *Phys. Rev. Lett.* **111**, 128701 (2013).
- [41] Magnani, M., Micenková, B. & Rossi, L. Combinatorial analysis of multiple networks. *arXiv preprint arXiv:1303.4986* (2013).
- [42] Coleman, J., Katz, E. & Menzel, H. The diffusion of an innovation among physicians. *Sociometry* **20**, 253–270 (1957).
- [43] Watts, D. J. & Strogatz, S. H. Collective dynamics of small-world networks. *Nature* **393**, 440–442 (1998).
- [44] Erdos, P. & Rényi, A. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci* **5**, 17–60 (1960).
- [45] Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).
- [46] Li, J.-W., Shen, Z.-S., Wang, W.-X., Grebogi, C. & Lai, Y.-C. Universal data-based method for reconstructing complex networks with binary-state dynamics. *Phys. Rev. E* **95**, 032303 (2017).