Network Reconstruction and Community Detection from Dynamics

Tiago P. Peixoto[1,2,3,*]

[1]*Department of Network and Data Science, Central European University, H-1051 Budapest, Hungary*
[2]*ISI Foundation, Via Chisola 5, 10126 Torino, Italy*
[3]*Department of Mathematical Sciences, University of Bath, Claverton Down, Bath BA2 7AY, United Kingdom*

We present a scalable nonparametric Bayesian method to perform network reconstruction from observed functional behavior that at the same time infers the communities present in the network. We show that the joint reconstruction with community detection has a synergistic effect, where the edge correlations used to inform the existence of communities are also inherently used to improve the accuracy of the reconstruction which, in turn, can better inform the uncovering of communities. We illustrate the use of our method with observations arising from epidemic models and the Ising model, both on synthetic and empirical networks, as well as on data containing only functional information.

The observed functional behavior of a wide variety large-scale system is often the result of a network of pairwise interactions. However, in many cases, these interactions are hidden from us, either because they are impossible to measure directly, or because their measurement can be done only at significant experimental cost. Examples include the mechanisms of gene and metabolic regulation [1], brain connectivity [2], the spread of epidemics [3], systemic risk in financial institutions [4], and influence in social media [5]. In such situations, we are required to infer the network of interactions from the observed functional behavior. Researchers have approached this reconstruction task from a variety of angles, resulting in many different methods, including thresholding the correlation between time series [6], inversion of deterministic dynamics [7–9], statistical inference of graphical models [10–14] and of models of epidemic spreading [15–20], as well as approaches that avoid explicit modeling, such as those based on transfer entropy [21], Granger causality [22], compressed sensing [23–25], generalized linearization [26], and matching of pairwise correlations [27,28].

In this Letter, we approach the problem of network reconstruction in a manner that is different from the aforementioned methods in two important ways. First, we employ a nonparametric Bayesian formulation of the problem, which yields a full posterior distribution of possible networks that are compatible with the observed dynamical behavior. Second, we perform network reconstruction jointly with community detection [29], where, at the same time as we infer the edges of the underlying network, we also infer its modular structure [30]. As we will show, while network reconstruction and community detection are desirable goals on their own, joining these two tasks has a synergistic effect, whereby the detection of communities significantly increases the accuracy of the reconstruction, which in turn improves the discovery of the communities, when compared to performing these tasks in isolation.

Some other approaches combine community detection with functional observation. Berthet *et al.* [31] derived necessary conditions for the exact recovery of group assignments for dense weighted networks generated with community structure given observed microstates of an Ising model. Hoffmann *et al.* [32] proposed a method to infer community structure from time-series data that bypasses network reconstruction by employing a direct modeling of the dynamics given the group assignments, instead. However, neither of these approaches attempt to perform network reconstruction together with community detection. Furthermore, they are tied down to one particular inverse problem, and as we will show, our general approach can be easily extended to an open-ended variety of functional models.

Bayesian network reconstruction.—We approach the network reconstruction task similarly to the situation where the network edges are measured directly, but via an uncertain process [33,34]: If $\mathcal{D}$ is the measurement of some process that takes place on a network, we can define a posterior distribution for the underlying adjacency matrix A via Bayes' rule

$$P(A|\mathcal{D}) = \frac{P(\mathcal{D}|A)P(A)}{P(\mathcal{D})}, \tag{1}$$

where $P(\mathcal{D}|A)$ is an arbitrary forward model for the dynamics given the network, $P(A)$ is the prior information on the network structure, and $P(\mathcal{D}) = \sum_A P(\mathcal{D}|A)P(A)$ is a normalization constant comprising the total evidence for the data $\mathcal{D}$. We can unite reconstruction with community detection via an, at first, seemingly minor, but ultimately consequential modification of the above equation where we

introduce a structured prior $P(A|b)$ where b represents the partition of the network in communities, i.e., $b = \{b_i\}$, where $b_i \in \{1, ..., B\}$ is group membership of node i. This partition is unknown, and is inferred together with the network itself, via the joint posterior distribution

$$P(A, b|\mathcal{D}) = \frac{P(\mathcal{D}|A)P(A|b)P(b)}{P(\mathcal{D})}. \quad (2)$$

The prior $P(A|b)$ is an assumed generative model for the network structure. In our work, we will use the degree-corrected stochastic block model (DC-SBM) [35], which assumes that, besides differences in degree, nodes belonging to the same group have statistically equivalent connection patterns, according to the joint probability

$$P(A|\lambda, \kappa, b) = \prod_{i<j} \frac{e^{-\kappa_i \kappa_j \lambda_{b_i, b_j}} (\kappa_i \kappa_j \lambda_{b_i, b_j})^{A_{ij}}}{A_{ij}!}, \quad (3)$$

with λ_{rs} determining the average number of edges between groups r and s and κ_i the average degree of node i. The marginal prior is obtained by integrating over all remaining parameters weighted by their respective prior distributions,

$$P(A|b) = \int P(A|\lambda, \kappa, b)P(\kappa|b)P(\lambda|b)d\kappa d\lambda, \quad (4)$$

which can be computed exactly for standard prior choices, although it can be modified to include hierarchical priors that have an improved explanatory power [36] (see Supplemental Material [37] for a concise summary.).

The use of the DC-SBM as a prior probability in Eq. (2) is motivated by its ability to inform link prediction in networks where some fraction of edges have not been observed or have been observed erroneously [34,39]. The latent conditional probabilities of edges existing between groups of nodes is learned by the collective observation of many similar edges, and these correlations are leveraged to extrapolate the existence of missing or spurious ones. The same mechanism is expected to aid the reconstruction task, where edges are not observed directly, but the observed functional behavior yields a posterior distribution on them, allowing the same kind of correlations to be used as an additional source of evidence for the reconstruction, going beyond what the dynamics alone says.

Our reconstruction approach is finalized by defining an appropriate model for the functional behavior, determining $P(\mathcal{D}|A)$. Here, we will consider two kinds of indirect data. The first comes from a susceptible-infected-susceptible (SIS) epidemic spreading model [40], where $\sigma_i(t) = 1$ means node i is infected at time t, 0, otherwise. The likelihood for this model is

$$P(\sigma|A, \tau, \gamma) = \prod_t \prod_i P(\sigma_i(t)|\sigma(t-1)), \quad (5)$$

where

$$P(\sigma_i(t)|\sigma(t-1))$$
$$= f(e^{m_i(t-1)}, \sigma_i(t))^{1-\sigma_i(t-1)} \times f(\gamma, \sigma_i(t))^{\sigma_i(t-1)} \quad (6)$$

is the transition probability for node i at time t, with $f(p, \sigma) = (1-p)^\sigma p^{1-\sigma}$, and where $m_i(t) = \sum_j A_{ij} \ln(1 - \tau_{ij})\sigma_j(t)$ is the contribution from all neighbors of node i to its infection probability at time t. In the equations above, the value τ_{ij} is the probability of an infection via an existing edge (i, j), and γ is the $1 \rightarrow 0$ recovery probability. With these additional parameters, the full posterior distribution for the reconstruction becomes

$$P(A, b, \tau|\sigma) = \frac{P(\sigma|A, \tau, \gamma)P(A|b)P(b)P(\tau)}{P(\sigma|\gamma)}. \quad (7)$$

Since $\tau_{ij} \in [0, 1]$, we use the uniform prior $P(\tau) = 1$. Note, also, that the recovery probability γ plays no role on the reconstruction algorithm, since its term in the likelihood does not involve A [and, hence, gets cancelled out in the denominator $P(\sigma|\gamma) = P(\gamma|\sigma)P(\sigma)/P(\gamma)$]. This means that the above posterior only depends on the infection events $0 \rightarrow 1$ and, thus, is also valid without any modifications to all epidemic variants susceptible-infected (SI), susceptible-infected-recovered (SIR), susceptible-exposed-infected-recovered (SEIR), etc., [40], since the infection events occur with the same probability for all these models.

The second functional model we consider is the Ising model, where spin variables on the nodes $s \in \{-1, 1\}^N$ are sampled according to the joint distribution

$$P(s|A, \beta, J, h) = \frac{\exp\left(\beta \sum_{i<j} J_{ij} A_{ij} s_i s_j + \sum_i h_i s_i\right)}{Z(A, \beta, J, h)}, \quad (8)$$

where β is the inverse temperature, J_{ij} is the coupling on edge (i, j), h_i is a local field on node i, and $Z(A, \beta, J, h) = \sum_s \exp(\beta \sum_{i<j} J_{ij} A_{ij} s_i s_j + \sum_i h_i s_i)$ is the partition function. Note that this is not a dynamical model as each microstate s is sampled independently according to the above distribution. Unlike the SIS model considered before, this distribution cannot be written in closed form since $Z(A, \beta, J, h)$ cannot be computed exactly, rendering the reconstruction problem intractable. Therefore, instead, we make use of the pseudolikelihood approximation [41], which is very accurate for the purpose at hand [14], where we approximate Eq. (8) as a product of (properly normalized) conditional probabilities for each spin variable s_i

$$P(s|A, \beta, J, h) = \prod_i \frac{\exp(\beta s_i \sum_j J_{ij} A_{ij} s_j + h_i s_i)}{2\cosh(\beta \sum_j J_{ij} A_{ij} s_j + h_i)}. \quad (9)$$

With the above likelihood, reconstruction is performed by observing a set of M microstates $\bar{s} = \{s_1, ..., s_M\}$, sampled

according to $P(\bar{s}|A,\beta,J,h) = \prod_l P(s_l|A,\beta,J,h)$, which yields the posterior distribution

$$P(A,b,\beta,J,h|\bar{s})$$
$$= \frac{P(\bar{s}|A,\beta,J,h)P(\beta)P(h)P(J|A)P(A|b)P(b)}{P(\bar{s})}. \quad (10)$$

In the above, we use uniform priors $P(J|A) = \prod_{ij}[-1/2 < J_{ij} < 1/2]^{A_{ij}}$, thus, forcing, without loss of generality, the values of J_{ij} to lie in the shifted unit interval $[-1/2, 1/2]$. For the remaining parameters, we use uniform priors, $P(h) \propto 1$ and $P(\beta) \propto 1$, for $\beta \in [-\infty, \infty]$ and $h \in [-\infty, \infty]^N$.

For any of the above posterior distributions, we perform sampling using a Markov chain Monte Carlo procedure: For each proposal $A \to A'$, it is accepted with the Metropolis-Hastings probability [42,43]

$$\min\left(1, \frac{P(A',b,\theta|\mathcal{D})}{P(A,b,\theta|\mathcal{D})}\frac{P(A' \to A)}{P(A \to A')}\right),$$

and likewise, for the node partition $b \to b'$, and any of the remaining parameters $\theta \to \theta'$. Note that the acceptance probability does not require the intractable normalization constant $P(\mathcal{D})$ to be computed. For both functional models considered, a whole sweep over E entries of the adjacency matrix and N nodes is done in time $O(EM + N\langle k \rangle)$, where M is the number of data samples per node, allowing the method to be applied for large systems. We summarize and give more details about the technical aspects of the algorithm in the Supplemental Material [37].

Synthetic networks.—We begin by investigating the reconstruction performance of networks sampled from the planted partition (PP) model, i.e., a DC-SBM with $\kappa_i = 1$, $\lambda_{rs} = \lambda_{\text{in}}\delta_{rs} + \lambda_{\text{out}}(1 - \delta_{rs})$, with $\lambda_{\text{in}} = \langle k \rangle[1 + \epsilon(B-1)]/N$ and $\lambda_{\text{out}} = \langle k \rangle(1 - \epsilon)/N$, where $\epsilon = N(\lambda_{\text{in}} - \lambda_{\text{out}})/\langle k \rangle B$ controls the strength of the modular structure. The detectability threshold for this model is given by $\epsilon^* = 1/\sqrt{\langle k \rangle}$, below which it is impossible to recover the planted community structure [44]. Given a network A^* from this model, we sample M independent Ising microstates s according to Eq. (8) using $J_{ij} = 1$, $h_i = 0$, and $\beta = \beta^*$ being the critical inverse temperature for the particular network. We compare two inference approaches: In the first, we sample both the reconstructed network as well as its community structure form the joint posterior of Eq. (10). In the second approach, we perform reconstruction and community detection separately, by first performing reconstruction in isolation, by replacing the DC-SBM prior $P(A|b)$ by the likelihood of an Erdős-Rényi model. We evaluate the quality of the reconstruction via the posterior similarity $S \in [0,1]$, defined as
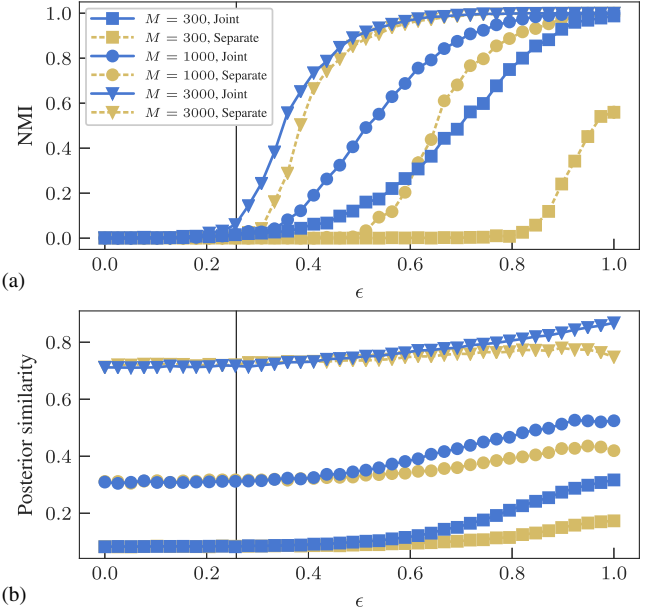

FIG. 1. Comparison between joint and separate reconstruction with community detection for a PP model with $N = 1000$, $\langle k \rangle = 15$, and $B = 10$. (a) Normalized mutual information (NMI) between inferred and planted node partitions, as a function of the model parameter ϵ, for several values of the number of samples M from the Ising model described in the text. (b) Posterior similarity between planted and inferred networks, for the same cases as in (a). The vertical line marks the detectability threshold $\epsilon = 1/\sqrt{\langle k \rangle}$.

$$S(A^*,\boldsymbol{\pi}) = 1 - \frac{\sum_{i<j}|A^*_{ij} - \pi_{ij}|}{\sum_{i<j}|A^*_{ij} + \pi_{ij}|}, \quad (11)$$

where A^* is the true network and $\boldsymbol{\pi}$ is the marginal posterior probability for each edge, i.e., $\pi_{ij} = \sum_{A,b,\theta} A_{ij}P(A,b,\theta|\mathcal{D})$. A value $S = 1$ means perfect reconstruction. We then perform community detection *a posteriori* by obtaining the maximum marginal point estimate

$$\hat{A}_{ij} = \begin{cases} 1 & \text{if } \pi_{ij} > 1/2, \\ 0 & \text{if } \pi_{ij} < 1/2, \end{cases} \quad (12)$$

and then sampling from the posterior $P(b|\hat{A})$. Figure 1 contains the comparison between both approaches for networks sampled from the PP model, which shows how sampling from the joint posterior improves both the reconstruction as well as community detection. For the latter, the joint inference allows the detection all the way down to the detectability threshold, for the examples considered, which, otherwise, is not possible with the separate method.

Real networks with synthetic dynamics.—Now, we investigate the reconstruction of networks not generated by the DC-SBM. We take two empirical networks, the
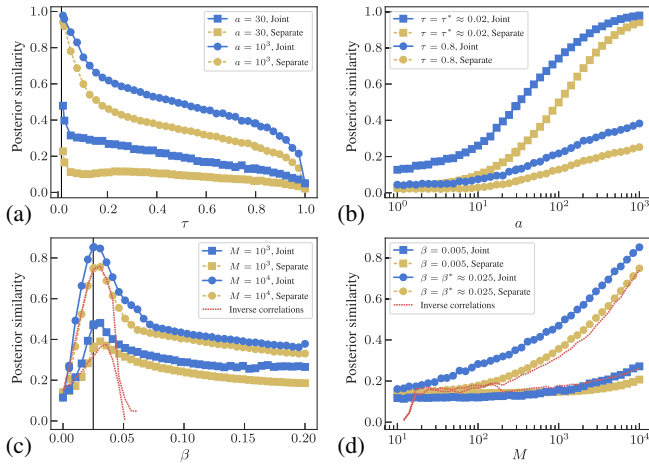
FIG. 2. Reconstruction results for simulated dynamics on empirical networks, comparing separate and joint reconstruction with community detection. (a) and (b) correspond to a SIS dynamics on global airport data, using $\tau_{ij} = \tau$, $\gamma = 1$, for different values of the infection probability τ and node activity a (defined as the number of infection events per node), and (c) and (d) the Ising model on a food web, using $J_{ij} = 1$ and $h_i = 0$. The dashed red line corresponds to the inverse correlation method for the Ising model. The solid vertical line marks the critical value for each model.
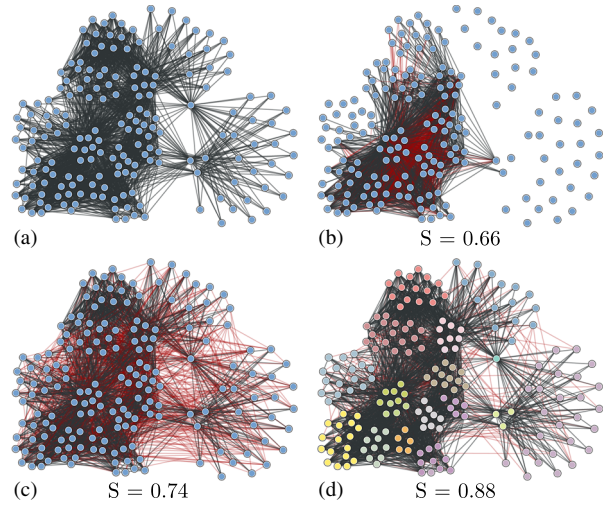

FIG. 3. Reconstruction of a food web network [46] from $M = 10^4$ samples of an Ising model at critical temperature. Edges marked in red are erroneous in the reconstruction. (a) Original network. (b) Optimal correlation thresholding. (c) Inverse correlations. (d) Joint reconstruction with community detection. The legends show the values of the posterior similarity [Eq. (11)].

worldwide network of $N = 3286$ airports [45] with $E = 39430$ edges, and a food web from Little Rock Lake [46], containing $N = 183$ nodes and $E = 2434$ edges, and we sample from the SIS (mimicking the spread of a pandemic) and Ising model (representing simplified interspecies interactions) on them, respectively, and evaluate the reconstruction obtained via the joint and separate inference with community detection, with results shown in Fig. 2. As is also the case for synthetic networks, the reconstruction quality is significantly improved by performing joint community detection [47]. The quality of the reconstruction peaks at the critical threshold for each model, at which the sensitivity to perturbations is the largest. As the number of observed samples increases, so does the quality of the reconstruction, and the relative advantage of the joint reconstruction diminishes, as the data eventually "washes out" the contribution from the prior. For the Ising model, we compare the results of our method with the mean-field inverse correlations method [14], i.e., $\beta A_{ij} J_{ij} = [C^{-1}]_{ij}$, where $C_{ij} = \langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle$ is the connected correlation matrix. As seen in Fig. 2, this simpler reconstruction method can be just as accurate as our separate reconstruction approach, but only close to the critical point. For higher inverse temperatures, the reconstruction deteriorates rapidly and breaks down completely as the system becomes locally magnetized, with whole rows and columns of the matrix C being equal to zero, causing it to be singular [48]. In such situations, this kind of approach requires explicit regularization techniques [49], which become

unnecessary with our Bayesian method. The joint inference with community structure improves the reconstruction even further, beyond what is obtainable with typical inverse Ising methods, since it incorporates a different source of evidence.

In Fig. 3, we show a comparison of the reconstruction of the food web network from a simulated Ising model, using different approaches. Optimal thresholding corresponds to the naive approach of imputing the existence of an edge to the connected correlation between two nodes exceeding a threshold c^*, i.e., $\pi_{ij} = \{1 \text{ if } C_{ij} > c^*, 0, \text{otherwise}\}$. The value of c^* was chosen to maximize the posterior similarity, which represents the best possible reconstruction achievable with this method. Nevertheless, the network thus obtained is severely distorted. The inverse correlation method comes much closer to the true network, but is superseded by the joint inference with community detection.

Empirical dynamics.—We turn to the reconstruction from observed empirical dynamics with unknown underlying interactions. The first example is the sequence of $M = 619$ votes of $N = 575$ deputies in the 2007 to 2011 session of the lower chamber of the Brazilian congress. Each deputy voted yes, no, or abstained for each legislation, which we represent as $\{1, -1, 0\}$, respectively. Since the temporal ordering of the voting sessions is likely to be of secondary importance to the voting outcomes, we assume the votes are sampled from an Ising model [the addition of zero-valued spins changes Eq. (9) only slightly by replacing $2\cosh(x) \to 1 + 2\cosh(x)$]. Figure 4 shows the result of the reconstruction, where the division of the nodes uncovers a cohesive government and a split
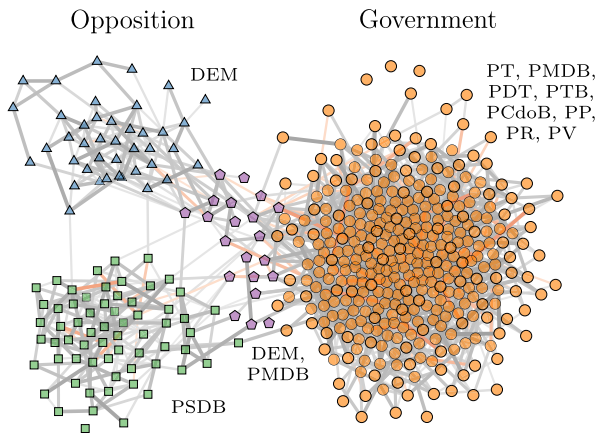
FIG. 4. Reconstruction of the interactions between members of the lower house of the Brazilian congress from the voting patterns of the 2007–2011 session, according to the Ising model. The node colors indicate the inferred groups. The edge thickness shows the posterior probability for each edge, and the color shows the magnitude of the coupling J_{ij}. The labels show the most frequent party membership for each group.

opposition, as well as a marginal center group, which correlates very well with the known party memberships and can be used to predict unseen voting behavior (see Supplemental Material [37] for more details). In Fig. 5, we show the result of the reconstruction of the directed network of influence between $N = 1833$ twitter users from 58224 retweets [50] using a SI epidemic model (the act of "retweeting" is modeled as an infection event,
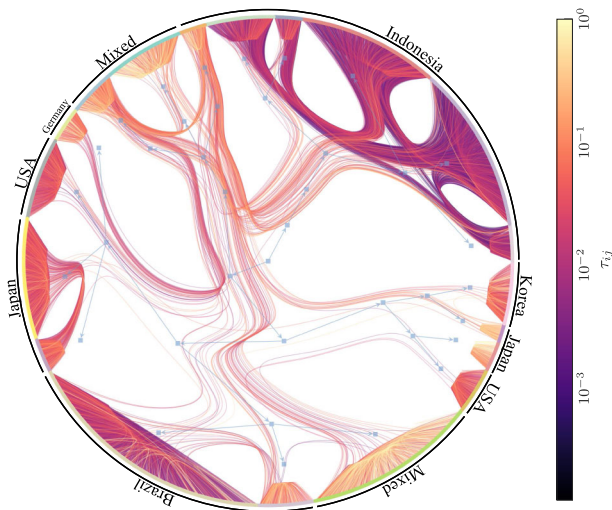

FIG. 5. Reconstruction of the directed network of influence between $N = 1833$ twitter users from 58224 retweets, using a SI infection model. The hierarchical division represents the inferred fit of the nested DC-SBM (see Refs. [51,52] for details on the layout algorithm), and the edge colors indicate the infection probabilities τ_{ij} as shown in the legend. The text labels show the dominating country membership for the users in each group.

using Eqs. (5) and (6) with $\gamma = 0$) and the nested DC-SBM. The reconstruction uncovers isolated groups with varying propensities to retweet, as well as groups that tend to influence a large fraction of users. By inspecting the geolocation metadata on the users, we see that the inferred groups amount, to a large extent, to different countries, although clear subdivisions indicate that this is not the only factor governing the influence among users (see Supplemental Material [37] for more details).

Conclusion.—We have presented a scalable Bayesian method to reconstruct networks from functional observations that uses the SBM as a structured prior and, hence, performs community detection together with reconstruction. The method is nonparametric and, hence, requires no prior stipulation of aspects of the network and size of the model, such as number of groups. By leveraging inferred correlations between edges, the SBM includes an additional source of evidence and, thereby, improves the reconstruction accuracy, which in turn also increases the accuracy of the inferred communities. The overall approach is general, requiring only appropriate functional model specifications, and can be coupled with an open ended variety of such models other than those considered here.

*peixotot@ceu.edu
[1] Y. Wang, T. Joshi, X.-S. Zhang, D. Xu, and L. Chen, Inferring gene regulatory networks from multiple microarray datasets, Bioinformatics **22**, 2413 (2006).

[2] M. Breakspear, Dynamic models of large-scale brain activity, Nat. Neurosci. **20**, 340 (2017).

[3] M. J. Keeling and P. Rohani, Estimating spatial coupling in epidemiological systems: a mechanistic approach, Ecol. Lett. **5**, 20 (2002).

[4] N. Musmeci, S. Battiston, G. Caldarelli, M. Puliga, and A. Gabrielli, Bootstrapping topological properties and systemic risk of complex networks using the fitness model, J. Stat. Phys. **151**, 720 (2013).

[5] E. Bakshy, I. Rosenn, C. Marlow, and L. Adamic, The role of social networks in information diffusion, in *Proceedings of the 21st International Conference on World Wide Web, WWW '12, Lyon, France* (ACM, New York, 2012), pp. 519–528.

[6] M. A. Kramer, U. T. Eden, S. S. Cash, and E. D. Kolaczyk, Network inference with confidence from multivariate time series, Phys. Rev. E **79**, 061916 (2009).

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

[8] S. G. Shandilya and M. Timme, Inferring network topology from complex dynamics, New J. Phys. **13**, 013004 (2011).

[9] M. Nitzan, J. Casadiego, and M. Timme, Revealing physical interaction networks from statistics of collective dynamics, Sci. Adv. **3**, e1600396 (2017).

[10] P. Abbeel, D. Koller, and A. Y. Ng, Learning factor graphs in polynomial time and sample complexity, J. Mach. Learn. Res. **7**, 1743 (2006).

[11] G. Bresler, E. Mossel, and A. Sly, Reconstruction of Markov random fields from samples: Some observations

and algorithms, in *Approximation, Randomization and Combinatorial Optimization. Algorithms and Techniques*, Lecture Notes in Computer Science (Springer, Berlin, Heidelberg, 2008), pp. 343–356.

[12] A. Montanari and J. A. Pereira, Which graphical models are difficult to learn, in *Advances in Neural Information Processing Systems*, edited by Y. Bengio, D. Schuurmans, J. D. Lafferty, C. K. I. Williams, and A. Culotta (Curran Associates, Inc., Red Hook, 2009), Vol. 22, pp. 1303–1311.

[13] H. Höfling and R. Tibshirani, Estimation of sparse binary pairwise Markov networks using pseudo-likelihoods, J. Mach. Learn. Res. **10**, 883 (2009).

[14] H. C. Nguyen, R. Zecchina, and J. Berg, Inverse statistical problems: From the inverse Ising problem to data science, Adv. Phys. **66**, 197 (2017).

[15] M. G. Rodriguez, J. Leskovec, and A. Krause, Inferring networks of diffusion and influence, in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '10* (ACM, New York, 2010), pp. 1019–1028.

[16] S. Myers and J. Leskovec, On the convexity of latent social network inference, in *Advances in Neural Information Processing Systems*, edited by J. D. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel, and A. Culotta (Curran Associates, Inc., Red Hook, 2010), Vol. 23, pp. 1741–1749.

[17] P. Netrapalli and S. Sanghavi, Learning the graph of epidemic cascades, in *Proceedings of the 12th ACM SIGMETRICS/PERFORMANCE Joint International Conference on Measurement and Modeling of Computer Systems, SIGMETRICS '12* (ACM, New York, 2012), pp. 211–222.

[18] C. Ma, H.-S. Chen, Y.-C. Lai, and H.-F. Zhang, Statistical inference approach to structural reconstruction of complex networks from binary time series, Phys. Rev. E **97**, 022301 (2018).

[19] B. Prasse and P. Van Mieghem, Maximum-likelihood network reconstruction for SIS processes is NP-hard, arXiv:1807.08630.

[20] B. Alfredo, I. Alessandro, and M. A. Paola, Network reconstruction from infection cascades, J. R. Soc. Interface **16**, 20180844 (2019).

[21] J. Runge, J. Heitzig, V. Petoukhov, and J. Kurths, Escaping the Curse of Dimensionality in Estimating Multivariate Transfer Entropy, Phys. Rev. Lett. **108**, 258701 (2012).

[22] J. Sun, D. Taylor, and E. M. Bollt, Causal network inference by optimal causation entropy, SIAM J. Appl. Dyn. Syst. **14**, 73 (2015).

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

[24] L. Ma, X. Han, Z. Shen, W.-X. Wang, and Z. Di, Efficient reconstruction of heterogeneous networks from time series via compressed sensing, PLoS One **10**, e0142837 (2015).

[25] X. Han, Z. Shen, W.-X. Wang, and Z. Di, Robust Reconstruction of Complex Networks from Sparse Data, Phys. Rev. Lett. **114**, 028701 (2015).

[26] J. Li, Z. Shen, W.-X. Wang, C. Grebogi, and Y.-C. Lai, Universal data-based method for reconstructing complex networks with binary-state dynamics, Phys. Rev. E **95**, 032303 (2017).

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

[28] P.-Y. Lai, Reconstructing network topology and coupling strengths in directed networks of discrete-time dynamics, Phys. Rev. E **95**, 022311 (2017).

[29] S. Fortunato and D. Hric, Community detection in networks: A user guide, Phys. Rep. **659**, 1 (2016).

[30] T. P. Peixoto, Bayesian stochastic blockmodeling, arXiv:1705.10225.

[31] Q. Berthet, P. Rigollet, and P. Srivastava, Exact recovery in the Ising blockmodel, arXiv:1612.03880.

[32] T. Hoffmann, L. Peel, R. Lambiotte, and N. S. Jones, Community detection in networks with unobserved edges, arXiv:1808.06079.

[33] M. E. J. Newman, Network structure from rich but noisy data, Nat. Phys. **14**, 542 (2018).

[34] T. P. Peixoto, Reconstructing Networks with Unknown and Heterogeneous Errors, Phys. Rev. X **8**, 041011 (2018).

[35] B. Karrer and M. E. J. Newman, Stochastic blockmodels and community structure in networks, Phys. Rev. E **83**, 016107 (2011).

[36] T. P. Peixoto, Nonparametric Bayesian inference of the microcanonical stochastic block model, Phys. Rev. E **95**, 012317 (2017).

[37] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.123.128301 for summary of the full generative model used, details of the inference algorithm and more information on the analysis of empirical data, which also contains Ref. [38].

[38] T. P. Peixoto, Efficient Monte Carlo and greedy heuristic for the inference of stochastic block models, Phys. Rev. E **89**, 012804 (2014).

[39] R. Guimerà and M. Sales-Pardo, Missing and spurious interactions and the reconstruction of complex networks, Proc. Natl. Acad. Sci. U.S.A. **106**, 22073 (2009).

[40] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, Epidemic processes in complex networks, Rev. Mod. Phys. **87**, 925 (2015).

[41] J. Besag, Spatial interaction and the statistical analysis of lattice systems, J. R. Stat. Soc. Ser. B **36**, 192 (1974).

[42] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, Equation of state calculations by fast computing machines, J. Chem. Phys. **21**, 1087 (1953).

[43] W. K. Hastings, Monte Carlo sampling methods using Markov chains and their applications, Biometrika **57**, 97 (1970).

[44] A. Decelle, F. Krzakala, C. Moore, and L. Zdeborová, Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications, Phys. Rev. E **84**, 066106 (2011).

[45] Retrieved from https://openflights.org/.

[46] N. D. Martinez, Artifacts or attributes? Effects of resolution on the little rock lake food web, Ecol. Monogr. **61**, 367 (1991).

[47] Note that, in this case, our method also exploits the heterogeneous degrees in the network via the DC-SBM,

which can, in principle, also aid the reconstruction in addition to the community structure itself.

[48] Refinements of this approach including Thouless-Anderson-Palmer (TAP) and Bethe-Peierls (BP) corrections [14] yield the same performance for this example.

[49] A. Decelle and F. Ricci-Tersenghi, Pseudolikelihood Decimation Algorithm Improving the Inference of the Interaction Network in a General Class of Ising Models, Phys. Rev. Lett. **112**, 070603 (2014).

[50] N. O. Hodas and K. Lerman, The simple rules of social contagion, Sci. Rep. **4**, 4343 (2015).

[51] T. P. Peixoto, Hierarchical Block Structures and High-Resolution Model Selection in Large Networks, Phys. Rev. X **4**, 011047 (2014).

[52] D. Holten, Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data, IEEE Trans. Visualization Comput. Graphics **12**, 741 (2006).