Letter

Reconstructing networks from simple and complex contagions

Nicholas W. Landry,[1,2,3,*] William Thompson,[1,†] Laurent Hébert-Dufresne,[1,4,‡] and Jean-Gabriel Young[1,2,§]
[1]*Vermont Complex Systems Center, University of Vermont, Burlington, Vermont 05405, USA*
[2]*Department of Mathematics and Statistics, University of Vermont, Burlington, Vermont 05405, USA*
[3]*Department of Biology, University of Virginia, Charlottesville, Virginia 22903, USA*
[4]*Department of Computer Science, University of Vermont, Burlington, Vermont 05405, USA*

(Received 7 May 2024; accepted 23 August 2024; published 7 October 2024)

Network scientists often use complex dynamic processes to describe network contagions, but tools for fitting contagion models typically assume simple dynamics. Here, we address this gap by developing a nonparametric method to reconstruct a network and dynamics from a series of node states, using a model that breaks the dichotomy between simple pairwise and complex neighborhood-based contagions. We then show that a network is more easily reconstructed when observed through the lens of complex contagions if it is dense or the dynamic saturates, and that simple contagions are better otherwise.

DOI: 10.1103/PhysRevE.110.L042301

Introduction. Contacts between individuals are how diseases, information, and transmissible social behavior spread through a population. Classical epidemic models treat these contact as well mixed [1], but we now have overwhelming evidence that the structure of these social systems can strongly influence the outcomes of a contagion—for example, the epidemic threshold [2], extent [3,4], or the time-scale of spread and the distribution of secondary infections [5]. Despite the strength of the evidence for structural models, network methods remain difficult to deploy beyond the theoretical laboratory, partly because detailed contact patterns are seldom observed and instead need to be inferred from messy and incomplete observational data [6,7]. In fact, this inference problem—network reconstruction—is still a fundamental problem in network science [6,8].

Recent progress suggests that state-of-the-art reconstruction is possible with Bayesian frameworks [9,10] in cases where the contagion spreads via single exposures to infectious individuals [9]. Other recent work has shown that network reconstruction is possible using measurements from repeated simple cascades [10], or binary state time series [11]. Less is known about the inference of social contagions, often thought to spread due to multiple exposures [12–14]. Previous studies have focused on distinguishing different types of contagions, whether interacting contagions and social reinforcement [15], simple and complex contagions [16,17], or heterogeneous and complex contagions [18], but these assume the underlying contact network is known and there is no emphasis on reconstruction itself.

In this Letter, we develop a nonparametric approach to network reconstruction from contagion data that does not assume the contagion dynamics are simple or complex. Using a time series of binary node states as input, we show how to jointly identify the network's structure and the dynamical rules that generated the contagion. Our method is conceptually related to that of Refs. [11,19,20], which all reconstruct a network and its dynamical rules from binary time-series. Unlike these previous studies, however, we put few restrictions on the contagion rules, and we generate a complete posterior distribution over parameters and contagion rules instead of point estimates.

We then use this framework to examine the difficulty of network reconstruction as a function of the contagion process's rules.

Neighborhood-based contagions on networks. We describe network contagions with a neighborhood-based susceptible–infected–susceptible (SIS) model [21], which encapsulates several simple and complex contagion processes as special cases. This contagion process is defined as follows: At time t, each infected nodes recover with probability γ and each susceptible nodes become infected with probability $c(\nu, \theta)$, where $c : \mathbb{N} \mapsto [0, 1]$ is the *contagion function*, ν is the number of infected neighbors of a node, and θ denotes any function parameters. We write this function as a contagion vector $\boldsymbol{c} = [c_0, \ldots, c_{N-1}]^T$ in nonparametric form, where c_ν denotes the probability of infection by ν infected neighbors and N is the number of nodes.

The classical network SIS and threshold models are special cases of this general model [21]. In discrete simple contagions, a node is infected with probability β from a single exposure, and exposures are modeled as independent, meaning that the contagion function is $c(\nu, \beta) = 1 - (1 - \beta)^\nu$. In complex contagions, multiple exposures are required for the process to spread, and they are most commonly studied using the threshold model [12,13], where a node adopts an opinion when the number of its neighbors holding that opinion exceeds a critical threshold τ. In that case, the contagion function becomes $c(\nu, \beta, \tau) = \beta \mathbf{1}_{\nu \geqslant \tau}$, where $\beta = 1$.

*Contact author: nicholas.landry@virginia.edu
†Contact author: william.thompson.2@uvm.edu
‡Contact author: laurent.hebert-dufresne@uvm.edu
§Contact author: jean-gabriel.young@uvm.edu

 ©2024 American Physical Society

This nonparametric model allows us to define a reconstruction process unrestricted by these two particular cases.

To describe the time evolution of this dynamical process mathematically, we denote the states of all nodes with a vector $\boldsymbol{x}(t) = [x_1(t), \ldots, x_N(t)]^T$, where $x_i(t)$ is the infection status of individual i at time t, with $x_i(t) = 0$ corresponding to a susceptible state and $x_i(t) = 1$ to the infected state. We track the states of nodes at unit time intervals $t = 0, 1, \ldots, T$. The collection of the state vectors is then the matrix $\boldsymbol{X} = [\boldsymbol{x}(0), \boldsymbol{x}(1), \ldots, \boldsymbol{x}(T)]$.

All the transition probabilities of the dynamics can be compactly expressed as

$$P(x_i(t+1) \mid x_i(t), \boldsymbol{A}, \theta) = b(\gamma, 1 - x_i(t+1))^{x_i(t)}\, b(c_{\nu_i(t)}, x_i(t+1))^{1-x_i(t)}, \quad (1)$$

where θ denotes all model parameters collectively—the recovery rate and the nonparametric contagion functions—and where $b(p, x) = p^x(1-p)^{1-x}$ is the Bernoulli probability mass function. In the absence of additional dependencies, we can write the joint probability of the transitions from step t to step $t+1$ as

$$P(\boldsymbol{x}(t+1) \mid \boldsymbol{x}(t), \boldsymbol{A}, \theta) = \prod_{i \in V} P(x_i(t+1) \mid x_i(t), \boldsymbol{A}, \theta). \quad (2)$$

The joint probability of a sequence of states $\boldsymbol{X}$ is then given as the product

$$P(\boldsymbol{X} = \boldsymbol{x} \mid \boldsymbol{A}, \theta) = \prod_{t=1}^{T-1} P(\boldsymbol{x}(t+1) \mid \boldsymbol{x}(t), \boldsymbol{A}, \theta), \quad (3)$$

since the dynamics are Markovian. By substituting Eq. (1) in Eq. (2) and then in Eq. (3) and rearranging terms, we can finally rewrite this probability as

$$P(\boldsymbol{X} \mid \boldsymbol{A}, \theta) = f(\gamma, \boldsymbol{X}) \prod_{\ell=0}^{N-1} c_\ell^{m_\ell} (1 - c_\ell)^{n_\ell}, \quad (4)$$

where

$$m_\ell = \sum_{t=0}^{T-1} \sum_{i \in V} [1 - x_i(t)] x_i(t+1) \mathbf{1}_{\nu_i(t) = \ell}, \quad (5a)$$

$$n_\ell = \sum_{t=0}^{T-1} \sum_{i \in V} [1 - x_i(t)][1 - x_i(t+1)] \mathbf{1}_{\nu_i(t) = \ell}, \quad (5b)$$

are the counts of infection events occurring and failing to occur when the number of infected neighbors is ℓ, and where $f(\gamma, \boldsymbol{X})$ is a function of γ and $\boldsymbol{X}$, but not of $\boldsymbol{A}$. [This factorization is due to Eq. (1), where $\nu_i(t)$ is the only term that depends on $\boldsymbol{A}$, implicitly.]

These equations determine the probability of a particular sequence of node states occurring for a given adjacency matrix $\boldsymbol{A}$ and set of dynamical parameters θ and, through Bayesian inference, they can be reversed to obtain the probability of these structural and dynamical properties given a sequence of system states. Figure 1 highlights this framework, which we now describe in more detail.

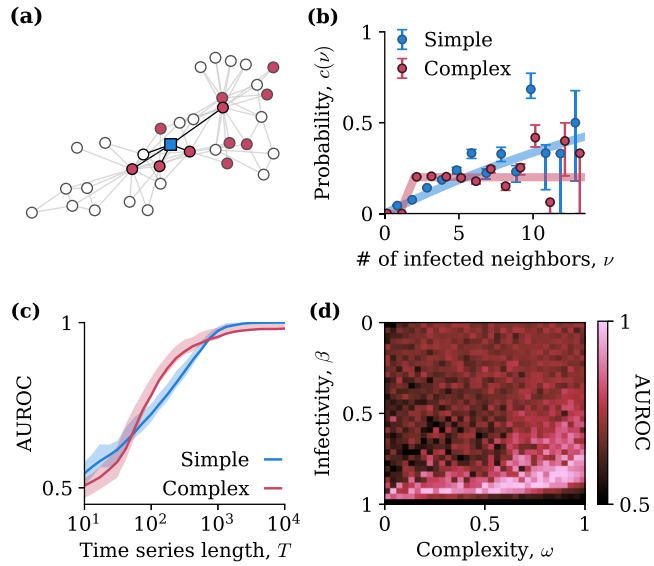

FIG. 1. Overview of the problem. (a) A contagion spreads on a network for T time steps, and we observe the resulting sequence of states $\boldsymbol{X}$. The probability that a susceptible node (white) becomes infected (red) at the next time step is a function $c(\nu)$ of the number of infected neighbors it has, e.g., $\nu = 4$ for the square node highlighted in blue. (b) We compute a nonparametric Bayesian estimate of the contagion function $c(\nu)$. Here, we show an estimate of $c(\nu)$ obtained from a single short realization of the dynamics when the network is known. Error bars show the 50% highest-density posterior interval (HDPI) of $c(\nu)$. (c) We estimate the network and the contagion function $c(\nu)$ simultaneously using the marginals of the posterior distribution, Eq. (7). The reconstruction error goes to 0 as the amount of data T goes to infinity. The shaded regions indicate the 50% HDPI, and lines show the median AUROC across 10^3 repetitions. (d) The reconstruction quality is determined by the shape of the contagion function, here demonstrated by varying its overall infectivity β and the level of complexity $\omega \in [0, 1]$. We use the parametrization $c(\nu, \beta, \omega) = (1-\omega)g + \omega h$, where $g(\nu, \beta) = 1 - (1-\beta)^\nu$ describes a simple contagion model, and $h(\nu, \beta) = \beta \mathbf{1}_{\nu \geqslant 2}$ describes a complex threshold model.

Network and dynamic reconstruction. Applying Bayes' rule, we write

$$P(\boldsymbol{A}, \theta, \rho \mid \boldsymbol{X}) = \frac{P(\boldsymbol{X} \mid \boldsymbol{A}, \theta)\, P(\boldsymbol{A} \mid \rho)\, P(\rho, \theta)}{P(\boldsymbol{X})}, \quad (6)$$

where $P(\boldsymbol{A} \mid \rho)$ is our prior over network structures $\boldsymbol{A}$, and where $\rho \in [0, 1]$ is a parameter of the network model, a simple Erdős-Rényi model of expected density ρ. This model assigns a probability $P(\boldsymbol{A} \mid \rho) = \rho^{|E|}(1-\rho)^{\binom{N}{2} - |E|}$ to networks with $|E|$ edges on N nodes. (But see Ref. [9] on replacing this simplistic assumption with more structured models.) Noticing that the parameters of the dynamical process $\theta = (\gamma, \boldsymbol{c})$ and the density ρ are all defined on the interval [0, 1], we choose independent (conjugate) beta priors with hyperparameters (a_γ, b_γ) for γ, and likewise for all the other parameters.

With these modeling choices, the posterior distribution can be written in closed form as

$$P(\boldsymbol{A}, \theta, \rho \mid \boldsymbol{X}) \propto f(\gamma, \boldsymbol{X}) \times \prod_{\ell=0}^{N-1} c_\ell^{m_\ell + a_\ell - 1}(1 - c_\ell)^{n_\ell + b_\ell - 1} \times \rho^{|E| + a_\rho - 1}(1-\rho)^{\binom{N}{2} - |E| + b_\rho - 1}, \tag{7}$$

where we have dropped normalization terms and redefined $f(\gamma, \boldsymbol{X})$ to include the prior density of γ also.

The network structure $\boldsymbol{A}$ is the primary target of network reconstruction, so we first focus on the marginal posterior distribution, which is obtained by integrating Eq. (7) over contagion vectors $\boldsymbol{c}$, recovery rates γ, and densities ρ. This yields

$$P(\boldsymbol{A} \mid \boldsymbol{X}) \propto \mathcal{B}\left(|E| + a_\rho, \binom{N}{2} - |E| + b_\rho\right) \times \prod_{\ell=0}^{N-1} \mathcal{B}(m_\ell + a_\ell, n_\ell + b_\ell), \tag{8}$$

where $\mathcal{B}(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the beta function. We generate samples from this distribution with a simple edge-flip Markov chain Monte Carlo (MCMC) algorithm [9]. We can then use these samples to compute a matrix $\boldsymbol{Q}$ of edge probability estimates, $\boldsymbol{Q} = \sum_{s=1}^{N_s} A^{(s)}/N_s$, where N_s is the number of samples. All our results are computed using chains starting from random initial conditions, with samples taken every 10^4 steps following an initial burn-in of 10^5 steps.

We augment these network samples with estimates of the dynamical parameters, $\theta = (\gamma, \boldsymbol{c})$, using the relation

$$P(\gamma, \boldsymbol{c}, \rho \mid \boldsymbol{X}, \boldsymbol{A}) \propto P(\gamma \mid \boldsymbol{X}, \boldsymbol{A}) P(\boldsymbol{c} \mid \boldsymbol{X}, \boldsymbol{A}) P(\rho \mid \boldsymbol{X}, \boldsymbol{A}),$$

which is obtained by conditioning Eq. (7) on $\boldsymbol{A}$, dropping all constants and noticing that the results factorize as a product of beta densities. More specifically, we find that $P(\gamma \mid \boldsymbol{X}, \boldsymbol{A})$ is a beta distribution with parameters $(h + a_\gamma, g + b_\gamma)$, where

$$g = \sum_{t=0}^{T-1} \sum_{i \in V} x_i(t) x_i(t+1),$$

$$h = \sum_{t=0}^{T-1} \sum_{i \in V} x_i(t)[1 - x_i(t+1)].$$

Similarly, the nonparametric infection probabilities $\boldsymbol{c}$ follow beta distributions $\boldsymbol{c}_\ell \sim \mathrm{Beta}(m_\ell + a_\ell, n_\ell + b_\ell)$. There is no need to sample ρ directly as $\boldsymbol{A}$ offers a good summary of structure; nonetheless, if needed, its posterior density is also a beta distribution.

Results. Figure 1 illustrates our nonparametric Bayesian approach using the Zachary's Karate Club network (henceforth ZKC) [22] as an example. We generate synthetic time series data comprising nodal states for $T = 10^3$ time steps with a recovery rate of $\gamma = 0.1$ starting from the initial condition $\boldsymbol{x}(0) = \mathbf{1}$ to prevent premature stochastic extinctions. We study the impact of the contagion function by running these simulations with both the SIS model, defined by $c(\nu, \beta) = 1 - (1-\beta)^\nu$, and the threshold contagion model $c(\nu, \beta, \tau) = \beta \mathbf{1}\nu \geqslant \tau$ with $\tau = 2$ (we also considered $\tau = 3$ in the Supplemental Material [23]). We find that imperfect reconstruction of the contagion function $\boldsymbol{c}$ is possible with short time series, as illustrated in Fig. 1(b). Values of ν for which we have little data—corresponding to high degrees—are more difficult to estimate accurately, as our nonparametric framework does not benefit from the strong inductive biases of simpler models.

Figure 1(c) describes the quality of the network reconstruction for ZKC, as measured by the Area Under the Receiver Operating Characteristic (AUROC) [24], as a function of the amount of time-series data. We calculate the AUROC by treating the edges $\boldsymbol{A}$ of the ZKC as binary labels and edge probabilities $\boldsymbol{Q}$ as posterior estimates for these labels. To make this comparison meaningful, we first select $\beta = 0.04$ for the simple contagion process and then match the maximum number of infection events per node of both processes in expectation by tuning the value β of the complex contagion process with the Robbins-Monro algorithm [25]. Both processes thus generate a similar *amount* of transition data, but their quality may vary.

We observe an intermediate value of T where complex contagions reconstruct the network more accurately than simple contagions. In addition, we find that increasing the amount of data leads to diminishing returns in reconstruction accuracy and that the accuracy reaches a plateau close to 1 for both dynamics in the limit of large T. Lastly, we construct a mixture of contagion functions parametrized by $\omega \in [0, 1]$ (where $\omega = 0$ corresponds to a simple contagion and $\omega = 1$ to a threshold contagion) Fig. 1(d) demonstrates that network reconstruction depends on both the infectivity of the contagion function and the complexity of the contagion.

Having confirmed that the method works well, we then investigate whether complex or simple contagion models lead to an easier network reconstruction problem. We explore this question through a comprehensive simulation experiment with generative network models that capture structural properties observed in empirical networks. The Erdős-Rényi model [27], which we use to examine the effect of network density, is characterized by p, the probability of two nodes connecting at random. We use the network configuration model (CM) [28] with a degree distribution $p(k) \propto k^\alpha$ supported on $[2, N-1]$ to observe the effects of degree heterogeneity, and we comment that when α is increased, the network density also increases. The clustered network model [29], which we use to represent clique structure, is constructed by forming a bipartite network, where type-1 nodes have degree 2 and type-2 nodes have degree s, and then projecting that network onto the type-1 nodes to create cliques of size s. The mean degree of the projected network scales with the square of the clique size. We fix the network size to $N = 50$ nodes and generated time series of $T = 2 \times 10^3$ steps. (We ran the study on two additional models; see the Supplemental Material [23] and Refs. [30–32] for more details.)

Figure 2 summarizes the result of this experiment using the *difference* in AUROC, with positive values (red) corresponding to regions of the parameter space where complex contagions outperform simple contagions. To complement this information, we also calculate the network density estimate quality, $\phi_\rho = 1 - \sum_{s=1}^{N_s} |\hat{\rho}_s - \rho|/N_s$ and compute a difference in performance $\Delta\phi_\rho$.

We find structural and dynamical regimes where different types of contagions outperform each other. The regions where

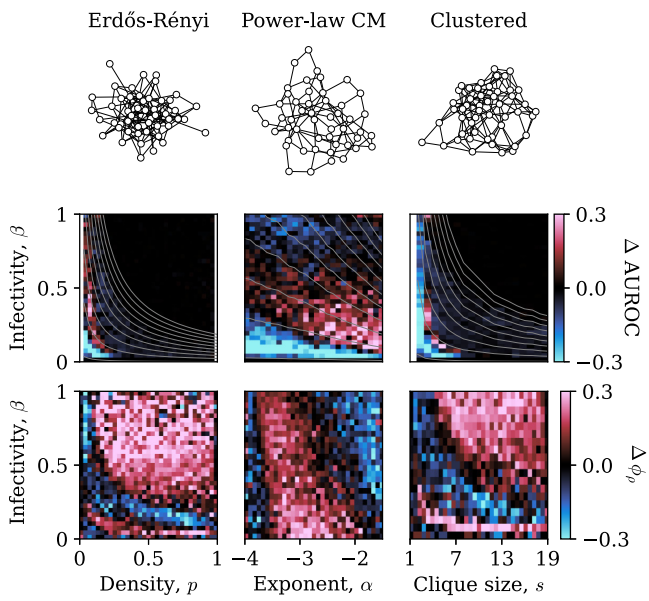

FIG. 2. The average difference in reconstruction performance between simple and complex contagions for the SIS model and the threshold model with $\tau = 2$. Red denotes regions where complex contagions outperform simple contagions. Blue denotes regions where simple contagions outperform complex contagions. The first row visualizes a sample from each network model; the second row compares the performance of the network reconstruction; and the third row compares the estimation of the network density. The gray lines show the basic reproduction number $R_0 = \beta\sigma(\boldsymbol{A})/\gamma \in \{1, 11, 21, ...\}$ [26], where $\sigma(\boldsymbol{A})$ is the spectral radius of $\boldsymbol{A}$, and grow when moving from the lower left corner to the upper right.

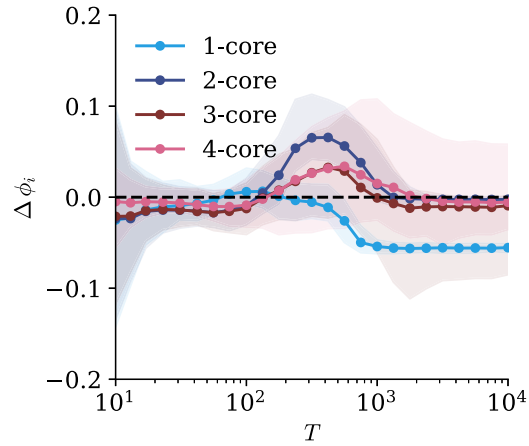

FIG. 3. Difference in reconstruction performance of the network SIS contagion process and the threshold contagion process with threshold $\tau = 2$, for nodes of various coreness and dynamical parameters identical to those chosen in Fig. 1(c). This illustrates that for intermediate amounts of infection data, complex contagions outperform simple contagions due to their recovery of $k \geqslant 2$-core nodes. The shaded regions represent the 50% HDPI, and the lines are the median of 1000 realizations.

simple contagions outperform complex contagions can be sufficiently explained by the basic reproduction number R_0, calculated as $\beta\sigma(\boldsymbol{A})/\gamma$, where $\sigma(\boldsymbol{A})$ is the spectral radius of $\boldsymbol{A}$. For all models, the simple contagion process outperforms the complex contagion process in a region roughly corresponding to $2 \leqslant R_0 \leqslant 6$ in the simple contagion. For higher R_0 values, the complex contagion process outperforms simple contagion as the infection time series saturates. Complex contagion never generates transitions with $\nu = 1$ since the threshold forbids it; this allows the algorithm to infer denser substructures without noise introduced by pairwise infection information. Just above this region, there is another region for large enough R_0 where simple contagion again outperforms complex contagion. This can be understood because complex contagion infers a network close to a complete network, whereas simple contagion infers an empty network—but with high confidence about several of the network links. For the densest of networks, these algorithms perform equally (notice that the densest power law CM is far from a complete network), because there is so much saturation that the elimination of pairwise information is not enough to overcome the confounding noise present when almost every node is infected at once.

There may be cases where, although our algorithm fails to place links correctly, it nonetheless can accurately recover the *number* of links, i.e., the density, ρ. Comparing the performance of complex contagions with respect to simple contagions in the bottom row of Fig. 2, we observe that these regions do not correspond to the performance differences with respect to AUROC. The regions in the upper right of the plots for the Erdős-Rényi and clustered network models where complex contagions vastly outperform simple contagions in estimating the network density are caused by bimodality in the distribution of estimated ρ values inferred from simple contagion time series. This is because our MCMC sampler struggles to sample in this region and gets stuck in local maxima due to the ruggedness of the likelihood landscape. In the lower left of the plots, where the variance in the estimated value of ρ is lower, we see similar behavior as described above for the AUROC.

A distinguishing feature of the threshold contagion is that it cannot spread to nodes within the 1-core. We observe this in Fig. 3, where we revisit the ZKC results to study the reconstruction performance of nodes with specific coreness values, calculated as the average of

$$\phi_i = 1 - \frac{1}{N_s}\sum_{s=1}^{N_s}\left(\frac{1}{N}\sum_{j=1}^{N}|\boldsymbol{Q}_{ij} - \boldsymbol{A}_{ij}|\right) \tag{9}$$

over all nodes i within the same coreness class. We find that nodes of coreness greater than one are responsible for complex contagions outperforming simple contagions in the region before AUROC plateaus [see Fig. 1(c)]. Similarly, as the amount of data is increased with T, we see that the differential performance of 1-core nodes seems to drive the improvement in the performance of simple contagions compared with complex contagions.

Discussion. By developing a nonparametric approach to the network inference problem, we avoid model-based biases and are able to study how the complexity of a dynamic process helps or hurts our ability to infer its rules and the network that supports it. We find that complex contagions can have an

advantage over equivalent simple contagions if they can avoid saturation and better resolve dense networks.

This result suggests that different dynamical processes have different statistical power to resolve different network structures, with threshold contagions being better in dense cliques when the dynamics are sufficiently intense. More generally, we could imagine optimizing or tuning a dynamic process to infer different network features. Future work should, therefore, explore other network structures and other types of contagions (e.g., relative thresholds, superlinear contagions). In doing so, we should be able to categorize different types of contagions not just by their global accuracy but by how well they can resolve different local features of networks.

More practically, it also suggests that different data streams allow us to learn different types of network features depending on the complexity of their generative mechanisms. Our results show that network reconstruction is much more feasible for diseases such as SARS-CoV-2 [33], Mpox [34], or rhinovirus in contrast to highly infectious diseases such as measles [35] or chickenpox [36]. In the context of social media where the spread of information is often modeled as a complex contagion process, our results imply that we will learn less than would be expected if we simply translate prior reconstructability results for simple contagion processes [9] with empirically determined estimates of R_0 [37]. For extremely viral trends or information, however, we may be able to recover the effective contact networks with more precision than previously estimated. These results will not only inspire future network reconstruction methods but also guide how well different types of experimental data can hope to inform network reconstruction.

All code and data used in this study are available on Github [38] and at Ref. [39]. All networks are visualized with XGI [40].

Acknowledgments. N.W.L., L.H.-D., and J.-G.Y. acknowledge support from the National Institutes of Health 1P20 GM125498-01 Centers of Biomedical Research Excellence Award. N.W.L. acknowledges support from the University of Virginia Prominence-to-Preeminence (P2PE) STEM Targeted Initiatives Fund, SIF176A Contagion Science.

[1] W. O. Kermack and A. G. McKendrick, A contribution to the mathematical theory of epidemics, Proc. R. Soc. Lond. A **115**, 700 (1927).

[2] R. Pastor-Satorras and A. Vespignani, Epidemic spreading in scale-free networks, Phys. Rev. Lett. **86**, 3200 (2001).

[3] J. C. Miller, A note on the derivation of epidemic final sizes, Bull. Math. Biol. **74**, 2125 (2012).

[4] L. Hébert-Dufresne, B. M. Althouse, S. V. Scarpino, and A. Allard, Beyond R_0: Heterogeneity in secondary infections and probabilistic epidemic forecasting, J. R. Soc. Interface. **17**, 20200393 (2020).

[5] B. M. Althouse, E. A. Wenger, J. C. Miller, S. V. Scarpino, A. Allard, L. Hébert-Dufresne, and H. Hu, Superspreading events in the transmission dynamics of SARS-CoV-2: Opportunities for interventions and control, PLoS Biol. **18**, e3000897 (2020).

[6] L. Peel, T. P. Peixoto, and M. De Domenico, Statistical inference links data and theory in network science, Nat. Commun. **13**, 6794 (2022).

[7] J.-G. Young, G. T. Cantwell, and M. E. J. Newman, Bayesian inference of network structure from unreliable data, J. Complex Netw. **8**, cnaa046 (2021).

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

[9] T. P. Peixoto, Network reconstruction and community detection from dynamics, Phys. Rev. Lett. **123**, 128301 (2019).

[10] C. Gray, L. Mitchell, and M. Roughan, Bayesian inference of network structure from information cascades, IEEE Trans. Signal Inf. Process. Netw. **6**, 371 (2020).

[11] 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).

[12] T. W. Valente, *Network Models of the Diffusion of Innovations* (Hampton Press, New York, 1995).

[13] D. J. Watts and P. S. Dodds, Influentials, networks, and public opinion formation, J. Consum. Res. **34**, 441 (2007).

[14] D. Centola, The spread of behavior in an online social network experiment, Science **329**, 1194 (2010).

[15] L. Hébert-Dufresne, S. V. Scarpino, and J.-G. Young, Macroscopic patterns of interacting contagions are indistinguishable from social reinforcement, Nat. Phys. **16**, 426 (2020).

[16] C. Murphy, E. Laurence, and A. Allard, Deep learning of contagion dynamics on complex networks, Nat. Commun. **12**, 4720 (2021).

[17] G. Cencetti, D. A. Contreras, M. Mancastroppa, and A. Barrat, Distinguishing simple and complex contagion processes on networks, Phys. Rev. Lett. **130**, 247401 (2023).

[18] G. St-Onge, L. Hébert-Dufresne, and A. Allard, Nonlinear bias toward complex contagion in uncertain transmission settings, Proc. Natl. Acad. Sci. USA **121**, e2312202121 (2024).

[19] Q.-M. Liu, C. Ma, B.-B. Xiang, H.-S. Chen, and H.-F. Zhang, Inferring network structure and estimating dynamical process from binary-state data via logistic regression, IEEE Trans. Syst. Man Cybern.: Syst. **51**, 4639 (2021).

[20] K. Liu, X. Lü, F. Gao, and J. Zhang, Expectation-maximizing network reconstruction and most applicable network types based on binary time series data, Physica D **454**, 133834 (2023).

[21] P. S. Dodds and D. J. Watts, Universal behavior in a generalized model of contagion, Phys. Rev. Lett. **92**, 218701 (2004).

[22] W. W. Zachary, An information flow model for conflict and fission in small groups, J. Anthropol. Res. **33**, 452 (1977).

[23] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.110.L042301 for additional figures and discussion.

[24] J. A. Hanley and B. J. McNeil, The meaning and use of the area under a receiver operating characteristic (ROC) curve, Radiology **143**, 29 (1982).

[25] H. Robbins and S. Monro, A stochastic approximation method, Ann. Math. Stat. **22**, 400 (1951).

[26] O. Diekmann, J. A. P. Heesterbeek, and J. A. J. Metz, On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations, J. Math. Biol. **28**, 365 (1990).

[27] P. Erdös and A. Rényi, On random graphs. I, Publ. Math. **6**, 290 (1959).

[28] B. K. Fosdick, D. B. Larremore, J. Nishimura, and J. Ugander, Configuring random graph models with fixed degree sequences, SIAM Rev. **60**, 315 (2018).

[29] M. E. J. Newman, Properties of highly clustered networks, Phys. Rev. E **68**, 026121 (2003).

[30] P. W. Holland, K. B. Laskey, and S. Leinhardt, Stochastic blockmodels: First steps, Social Networks **5**, 109 (1983).

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

[32] M. Jerrum and G. B. Sorkin, The Metropolis algorithm for graph bisection, Discret. Appl. Math. **82**, 155 (1998).

[33] H. Li, C.-J. Arcalas, J. Song, M. Rahmati, S. Park, A. Koyanagi, S. W. Lee, D. K. Yon, J. I. Shin, and L. Smith, Genetics, structure, transmission, epidemiology, immune response, and vaccine efficacies of the SARS-CoV-2 Delta variant: A comprehensive review, Rev. Med. Virol. **33**, e2408 (2023).

[34] G. Guzzetta, A. Mammone, F. Ferraro, A. Caraglia, A. Rapiti, V. Marziano, P. Poletti, D. Cereda, F. Vairo, G. Mattei, F. Maraglino, G. Rezza, and S. Merler, Early estimates of monkeypox incubation period, generation time, and reproduction number, Italy, May–June 2022, Emerg. Infect. Dis. **28**, 2078 (2022).

[35] F. M. Guerra, S. Bolotin, G. Lim, J. Heffernan, S. L. Deeks, Y. Li, and N. S. Crowcroft, The basic reproduction number (R_0) of measles: a systematic review, Lancet Infect. Dis. **17**, e420 (2017).

[36] T. Lee, J. Suh, J.-K. Choi, J. Lee, and S. H. Park, Estimating the basic reproductive number of varicella in South Korea incorporating social contact patterns and seroprevalence, Hum. Vaccin. Immunother. **17**, 2488 (2021).

[37] A. J. Kucharski, Modelling the transmission dynamics of online social contagion, arXiv:1602.00248.

[38] https://github.com/nwlandry/complex-network-reconstruction.

[39] N. Landry, W. Thompson, and J.-G. Young, nwlandry/complex-network-reconstruction (2024).

[40] N. W. Landry, M. Lucas, I. Iacopini, G. Petri, A. Schwarze, A. Patania, and L. Torres, XGI: A Python package for higher-order interaction networks, J. Open Source Softw. **8**, 5162 (2023).