

Message-passing approach for recurrent-state epidemic models on networks

Munik Shrestha

*Department of Physics and Astronomy, University of New Mexico, Albuquerque, New Mexico 87131, USA
and Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA*

Samuel V. Scarpino and Christopher Moore

Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA

(Received 7 May 2015; published 27 August 2015)

Epidemic processes are common out-of-equilibrium phenomena of broad interdisciplinary interest. Recently, dynamic message-passing (DMP) has been proposed as an efficient algorithm for simulating epidemic models on networks, and in particular for estimating the probability that a given node will become infectious at a particular time. To date, DMP has been applied exclusively to models with one-way state changes, as opposed to models like SIS and SIRS where nodes can return to previously inhabited states. Because many real-world epidemics can exhibit such recurrent dynamics, we propose a DMP algorithm for complex, recurrent epidemic models on networks. Our approach takes correlations between neighboring nodes into account while preventing causal signals from backtracking to their immediate source, and thus avoids “echo chamber effects” where a pair of adjacent nodes each amplify the probability that the other is infectious. We demonstrate that this approach well approximates results obtained from Monte Carlo simulation and that its accuracy is often superior to the pair approximation (which also takes second-order correlations into account). Moreover, our approach is more computationally efficient than the pair approximation, especially for complex epidemic models: the number of variables in our DMP approach grows as $2mk$ where m is the number of edges and k is the number of states, as opposed to mk^2 for the pair approximation. We suspect that the resulting reduction in computational effort, as well as the conceptual simplicity of DMP, will make it a useful tool in epidemic modeling, especially for high-dimensional inference tasks.

DOI: [10.1103/PhysRevE.92.022821](https://doi.org/10.1103/PhysRevE.92.022821)

PACS number(s): 64.60.aq, 89.20.-a

I. INTRODUCTION

Mathematical models of epidemic processes are intrinsically nonlinear and multiplicative. These models include the spread of disease [1,2], transmission of social behaviors [3–6], cascades of banking failures [7,8], forest fires [9–11], the propagation of marginal probabilities in constraint satisfaction problems [12,13], and the dynamics of magnetic and glassy systems [14].

The classical approach to modeling epidemics, such as the SIR model where each node is susceptible, infectious, or recovered, assumes that at any given time each individual exists in a single state or “compartment” [1,2]. To make these models analytically tractable, it is often assumed that the population is well mixed, so that interaction between any two individuals is equally likely; in physical terms, we assume the model is mean-field (also known as mass-action mixing in the epidemiology literature). Despite this unrealistic assumption, mean-field models capture some essential features of epidemics, such as a threshold above which we have an endemic phase with a nonzero fraction of infected individuals, and below which we have outbreaks of size $o(n)$ so that the equilibrium fraction of infected individuals is zero.

In reality, contacts between individuals in the population are often highly structured, with some pairs of individuals much more likely to interact than others due to location or demographics [6,15]. To relax the mean-field assumption, while retaining some measure of tractability, we can assume that individuals interact on a network whose structure captures the heterogeneity in the population [16,17]. However, replacing the mean-field approximation with a contact network substantially increases a model’s complexity.

One reasonable goal is to compute the one-point marginals, e.g., for each node i the probability $I_i(t)$ that i is infectious at time t . In addition to being of direct interest, these marginals help us perform tasks such as inferring the originator of an epidemic, determining an optimal set of nodes to immunize in order to minimize the final size of an outbreak, or calculating the probability that an entire group of nodes will remain uninfected after a fixed time [18–22].

We can always compute these marginals by performing Monte Carlo experiments. However, since we need to perform many independent trials in order to collect good statistics, this is computationally expensive on large networks. This problem is compounded if we need to scan through parameter space, or if we want to explore many different initial conditions, vaccination strategies, etc. Therefore, it would be desirable to compute these marginals using, say, a system of differential equations, with variables that directly model the probabilities of various events.

The most naive way to do this, as we review below, uses the one-point marginals themselves as variables. However, this approach completely ignores correlations between nodes. At the other extreme, to model the system exactly, we would need to keep track of the entire joint distribution: but if there are n individuals, each of which can be in one of k states, this results in a coupled system with k^n variables. This exponential scaling quickly renders most models computationally intractable, even on moderately sized networks.

In between these two extremes, we can approximate the joint distribution by “moment closure,” assuming that higher-order marginals can be written in terms of lower-order ones. This gives a hierarchy of increasingly accurate

(and computationally expensive) approximations, familiar in physics as cluster expansions. At the first level of this hierarchy we assume that the nodes are uncorrelated and approximate two-point marginals such as $[I_i(t) \wedge I_j(t)]$ (the probability that i and j are both infectious at time t) as $I_j(t)I_i(t)$. At the second level, commonly referred to in the epidemiology literature as the pair approximation, we close the hierarchy at the level of pairs $[I_i(t) \wedge I_j(t)]$ by assuming that three-point correlations can be factored in terms of two-point correlations. For a comprehensive review of these methods, see Refs. [17,23].

In this paper, we study an alternative method, namely *dynamic message-passing* (DMP). As in belief propagation [24,25], here variables or “messages” are defined on a network’s directed edges: for instance, $I_{j \rightarrow i}$ denotes the probability that j was infected by one of its neighbors other than i , so that the epidemic might spread from j to i . However, unlike belief propagation, where the posterior distributions are updated according to Bayes’ rule, here we write differential equations for the messages over time.

For many epidemic models, such as SI (susceptible-infectious), SIR (susceptible-infectious-recovered) and SEIR (susceptible-exposed-infectious-recovered), only one-way state changes can occur. For example, in the SIR model, once an individual has left the susceptible class and become infectious, they cannot return to being susceptible; once they become recovered, they are immune to future infections, and might as well be removed. For these nonrecurrent models, DMP is known to be an efficient algorithm to estimate $I_i(t)$, and it is exact on trees [26]; it can also be applied to threshold models [27–29] and used for inference [18].

However, for many real-world diseases individuals can return to previously inhabited states. In these *recurrent* models, such as SIS (susceptible-infectious-susceptible), SIRS (susceptible-infectious-recovered-susceptible), and SEIS (susceptible-exposed-infectious-susceptible), individuals can cycle through the states multiple times, often resulting in multiple waves of infection traveling through the population. The most obvious examples of recurrent models are seasonal influenza, where due to the evolution of the virus individuals are repeatedly infected during their lifetime [30], vaccination where protective immunity wanes over time [31], and diseases curable by treatment, which does not result in antibody-mediated immunity, such as gonorrhea [32]. In all three cases, individuals leave the susceptible class, only to return at some point in the future (although for influenza, it is worth mentioning that if the evolutionary rate of the virus is functionally related to the number of susceptible individuals, then the recovery rate may not be independent from the state of one’s neighbors). Unfortunately, the DMP approach of Ref. [26] cannot be directly extended to recurrent models, since their equations for messages only track the first time an individual makes the transition to a given state.

The purpose of this paper is to develop a novel DMP algorithm for recurrent models of epidemics on networks, which we call rDMP. We will show that rDMP gives very good approximations for marginal probabilities on networks and is often more accurate than the pair approximation. Moreover, whereas the pair approximation requires keeping track of mk^2 variables, if there are m edges and k states per node, rDMP requires just $2mk$ variables. For complex models

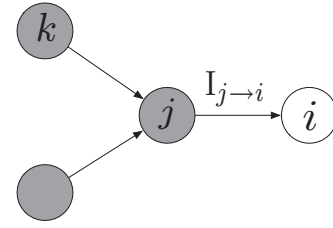


FIG. 1. We define messages on the directed edges of a network to carry causal information of the flow of contagion; e.g., $I_{j \rightarrow i}$ is the probability that j is infectious because it received the infection from a neighbor k other than i . This prevents signals from immediately backtracking to the node they came from and thus prevents what we term “echo chamber” infections.

where k is large—for instance, for diseases with multiple stages of infection or immunity, or multiple-disease epidemics where one disease makes individuals more susceptible to another one—rDMP provides a substantial reduction in the computational effort required. Finally, the rDMP approach is conceptually simple, making it easy to write down the system of differential equations for a wide variety of epidemic models.

II. MESSAGE-PASSING AND PREVENTING THE ECHO CHAMBER EFFECT

As shown in Fig. 1, the variables of rDMP are messages along directed edges of the network (in addition to one-point marginals). For instance, $I_{j \rightarrow i}$ is the probability that j is infectious because it was infected by one of its other neighbors k . The intuition behind this is the following, where we consider the SIS model as an example. If i is susceptible, the rate at which j will infect i is proportional to the probability I_j that j is infected. But when computing this rate, we only include the contribution to I_j that comes from neighbors other than i . In other words, we deliberately neglect the event that j receives the infection from i , and immediately passes it back to i , even if i has become susceptible in the intervening time.

This choice avoids a kind of “echo chamber” effect, where neighboring nodes artificially amplify each others’ probability of being infectious. For instance, consider a simple but pathological case of the SI model where there are only two nodes in the graph, i and j , with an edge between them as shown in Fig. 2. If the transmission rate is λ , and if we assume the nodes are independent (i.e., if we use first-order moment closure), we obtain the following differential equations:

$$\begin{aligned} \frac{dI_i}{dt} &= \lambda S_i I_j, \\ \frac{dI_j}{dt} &= \lambda S_j I_i, \end{aligned} \quad (1)$$

where $S_i(t) = 1 - I_i(t)$ and similarly for j .



FIG. 2. Two simple, yet illustrative, cases of networks, where the darker node is initially infectious. As we discuss, in these simple cases one can see the motivation for our approach to prevent infection signals from backtracking.

Now suppose that j is initially infectious with probability δ , and that i is initially susceptible; i.e., $I_j(0) = \delta$ and $I_i(0) = 0$. Since in the SI model nodes never recover, the infection will eventually spread from j to i , but only if i was infectious in the first place. Thus the marginals $I_i(t)$ and $I_j(t)$ should tend to δ as $t \rightarrow \infty$.

However, integrating Eq. (1) gives a different result. Once I_i becomes positive, dI_j/dt becomes positive as well, allowing i to infect j with the infection that it received from j in the first place. As a result, $I_j(t)$ approaches 1 as $t \rightarrow \infty$. Thus, the ‘‘echo chamber’’ between i and j leads to the absurd result that j eventually becomes infectious, even though with probability $1 - \delta$ there was no initial infection in the system.

In the rDMP approach, we fix this problem by replacing I_i and I_j with the messages they send each other,

$$\begin{aligned}\frac{dI_i}{dt} &= \lambda S_i I_{j \rightarrow i}, \\ \frac{dI_j}{dt} &= \lambda S_j I_{i \rightarrow j},\end{aligned}$$

so that i can only infect j if i received the infection from some node other than j . In this example, there are no other nodes, so if $I_{j \rightarrow i}(0) = \delta$ and $I_{i \rightarrow j}(0) = 0$, then $I_j(t) = \delta$ for all t as it should be.

Note that we do not claim that rDMP is exact in this case. In particular, as in Eq. (1), $I_i(t)$ tends to 1 as $t \rightarrow \infty$. This is because, unlike the system of Ref. [26], rDMP assumes that each time j infects i is independent from the previous.

In this two-node example, of course, the pair approximation is exact, since it maintains separate variables such as $[S_j \wedge I_k]$ for each of the joint states of the two nodes. However, the pair approximation is subject to other forms of the ‘‘echo chamber effect.’’ Consider a network with three nodes, as in Fig. 2 (right), where j is a common neighbor of i and k . The pair approximation assumes that, conditioned on the state of j , the states of i and k are independent; however, in a recurrent epidemic model, i and k could be correlated, for instance if j infected them both and then returned to the susceptible state. As a result, the pair approximation is vulnerable to a distance-two echo chamber, where i and k infect each other through j . As in the two-node case, rDMP prevents this.

Preventing backtracking completely may seem like a strong assumption, and in recurrent models it is *a priori* possible, for instance, for a node to reinfect the neighbor it was infected by. Despite the well-documented importance of recurrent infections for diseases including (but certainly not limited to) seasonal influenza [30], plasmodium malaria [33], and urinary tract infections [34], little is known about the source of recurrent infections. For certain sexually transmitted diseases such as gonorrhea [32] and repeated ringworm infections [35], there is evidence that backtracking plays a significant role; on the other hand, it may be that recurrent infections are caused by different strains, each of which is acting essentially without backtracking. Thus, while our nonbacktracking assumption is clearly invalid in some cases, we believe it is a reasonable approach for most recurrent state infections.

III. THE rDMP EQUATIONS FOR THE SIS, SIRS, AND SEIS MODELS

In this section, we illustrate the rDMP approach for several recurrent epidemic models. We start with the simplest, which is the SIS model where each node is either infectious (I) or susceptible (S). Infectious nodes infect their susceptible neighbors at rate λ , and their infections wane back into the susceptible state at rate ρ . We denote the probability that node i is infectious or susceptible by I_i and S_i , respectively. The objective then is to efficiently and accurately compute these probabilities as a function of time t .

We define variables or ‘‘messages’’ that live on the directed edges (i, j) of the network. The directed nature of these messages prevent infection from backtracking from an infectious node back to its infection source; e.g., if node i infects node j , then we prevent j from reinfecting i . In addition to tracking the one-point marginal I_j , we define a message $I_{j \rightarrow i}$ from j to i as the probability that j is in the infectious state as a result of being infected from one of its neighbors other than i . Given these incoming messages, the rate at which I_i evolves in time is given by

$$\frac{dI_i}{dt} = -\rho I_i + \lambda S_i \sum_{j \in \partial i} I_{j \rightarrow i}, \quad (2)$$

where ∂i denotes the neighbors of i . Similarly, the rate at which $I_{j \rightarrow i}$ evolves in time is given by

$$\frac{dI_{j \rightarrow i}}{dt} = -\rho I_{j \rightarrow i} + \lambda S_j \sum_{k \in \partial j \setminus i} I_{k \rightarrow j}, \quad (3)$$

where $k \in \partial j \setminus i$ denotes the neighbors of j excluding i .

For the SIRS model, we let ρ and γ denote the transition rates from infectious to recovered and from recovered to susceptible, respectively. Then the rDMP system for the SIRS model is given by

$$\frac{dI_{j \rightarrow i}}{dt} = -\rho I_{j \rightarrow i} + \lambda S_j \sum_{k \in \partial j \setminus i} I_{k \rightarrow j}, \quad (4)$$

which is coupled with the one-point marginals through

$$\begin{aligned}\frac{dS_i}{dt} &= \gamma R_i - \lambda S_i \sum_{j \in \partial i} I_{j \rightarrow i}, \\ \frac{dI_i}{dt} &= -\rho I_i + \lambda S_i \sum_{j \in \partial i} I_{j \rightarrow i},\end{aligned} \quad (5)$$

$$\frac{dR_i}{dt} = \rho I_i - \gamma R_i.$$

In the SEIS model, upon becoming exposed to an infected neighbor, susceptible nodes first go through a latent period called the exposed state. In this state, individuals are infected but not yet infectious. Exposed nodes become infectious at the rate ε , and infectious nodes again wane back to susceptible at rate ρ . The rDMP system for the SEIS model is

$$\begin{aligned}\frac{dE_{j \rightarrow i}}{dt} &= -\varepsilon E_{j \rightarrow i} + \lambda S_j \sum_{k \in \partial j \setminus i} I_{k \rightarrow j}, \\ \frac{dI_{j \rightarrow i}}{dt} &= -\rho I_{j \rightarrow i} + \varepsilon E_{j \rightarrow i},\end{aligned} \quad (6)$$

which is coupled with the one-point marginals as

$$\begin{aligned} \frac{dS_i}{dt} &= \rho I_i - \lambda S_i \sum_{j \in \partial i} I_{j \rightarrow i}, \\ \frac{dE_i}{dt} &= -\varepsilon E_i + \lambda S_i \sum_{j \in \partial i} I_{j \rightarrow i}, \\ \frac{dI_i}{dt} &= -\rho I_i + \varepsilon E_i. \end{aligned} \quad (7)$$

Note that here we track messages for the exposed state, in addition to one-point marginals, since they act as precursors for the infectious messages. There is no need to track messages for the susceptible state, since it does not cause state changes in its neighbors.

Generalizing these equations to more complex epidemic models with k different states, as opposed to three or four, is straightforward. Even in a model where every state can cause changes in a neighbor's state—for instance, where having susceptible neighbors speeds up the rate of recovery or where exposed nodes can also infect their neighbors at a lower rate—the total number of variables we need to track in a network with n nodes and m edges is at most $2mk$ in addition to the nk one-point marginals. In contrast, the pair approximation requires mk^2 states to keep track of the joint distribution of every neighboring pair.

IV. EXPERIMENTS IN REAL AND SYNTHETIC NETWORKS

In this section we report on numerical experiments for rDMP for the SIS and SIRS models on real and synthetic networks. As a performance metric, we use the average L_1 error per node between the marginals computed from rDMP and the true probabilities computed (up to sampling error)

using continuous-time Monte Carlo simulations. That is,

$$L_1^{\text{rDMP}}(t) = \frac{1}{n} \sum_i |I_i^{\text{MC}}(t) - I_i^{\text{rDMP}}(t)|, \quad (8)$$

We use this metric to compare the performance of rDMP with the independent-node approximation and the pair approximation, or equivalently first- and second-order moment closure [17,23]. As we will see, for a wide range of parameters, rDMP is more accurate than either of these approaches, even though it is more computationally efficient than the pair approximation.

In Fig. 3, we show results for the SIS model on Zachary's karate club [36]. On the left, we plot the marginal probability that a particular node is infectious as a function of time, estimated by rDMP and by first- and second-order moment closure, and compared to those estimates with the true marginals given by Monte Carlo simulation. On the right, we show the average L_1 error for the three methods. Here $\lambda = 0.1$, $\rho = 0.05$, and the initial condition consists of a single infected node (shown in red in the inset). The Monte Carlo results were averaged over 10^5 runs. We see that rDMP is significantly more accurate than the other two, except very early in the simulation.

As a further illustration, in Fig. 4 we show the steady-state marginal I_i for each node i (measured by running the system until $t = 50$, at which point $I_i(t)$ is nearly constant), with the same parameters and initial condition as described in the caption of Fig. 3. We show the true marginal of each node on the y axis and the marginals estimated by rDMP and the pair approximation on the x axis. If the estimated marginals were perfectly accurate, the points would fall on the line $y = x$. Both methods overestimate the marginals to some extent, but rDMP is more accurate than the pair approximation on every node. Thus, rDMP makes accurate estimates of the marginals on individual nodes, as opposed to just the average across the population.

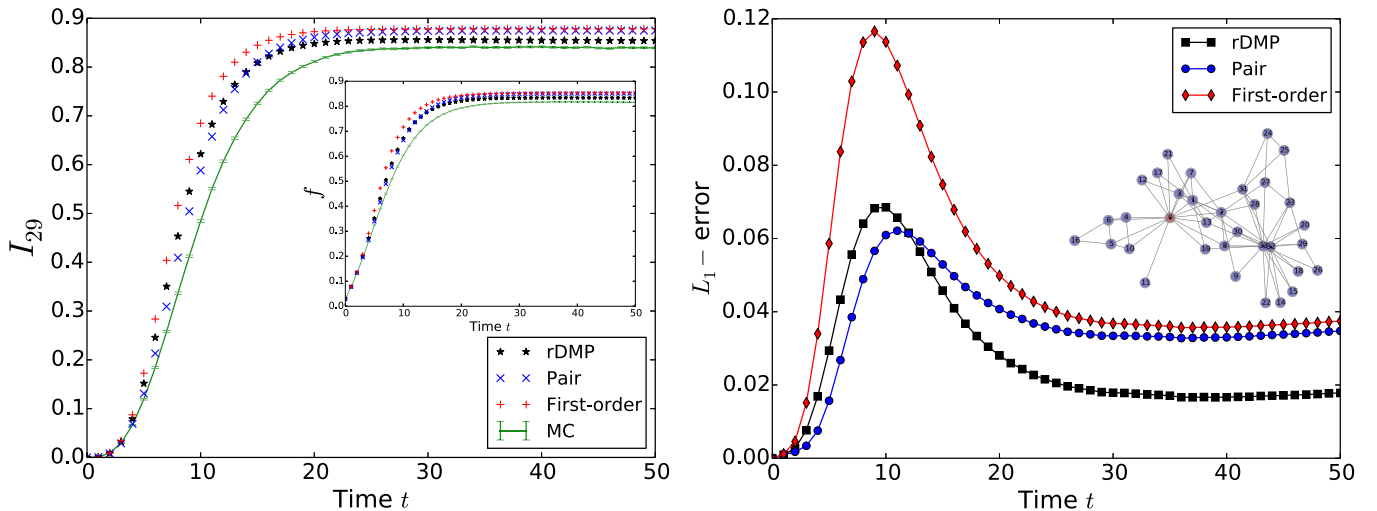


FIG. 3. (Color online) Results on the SIS model. On the left, the marginal probability that node 29 in Zachary's karate club (see inset on right) is infectious as a function of time. We compare the true marginal derived by 10^5 independent Monte Carlo simulations with that estimated by rDMP, the independent node approximation, and the pair approximation. In the inset, we show the fraction f of infectious nodes as a function of time. On the right is the L_1 error, averaged over all nodes; we see that rDMP is the most accurate of the three methods. Here the transmission rate is $\lambda = 0.1$, the waning rate is $\rho = 0.05$, and vertex 0 (colored red) was initially infected.

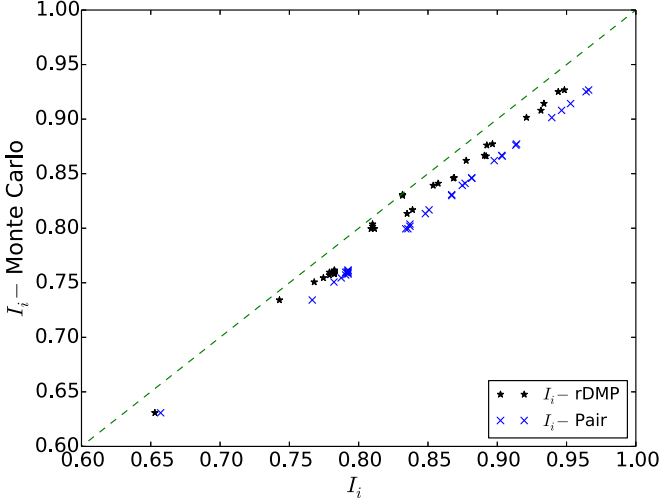


FIG. 4. (Color online) A scatter plot of the steady-state marginals I_i for the $n = 33$ nodes in Zachary's karate club, with the same parameters as described in the caption of Fig. 3. The vertical axis is the true marginal computed by Monte Carlo simulations; the horizontal axis is the estimated marginals from rDMP (black \star) and the pair approximation (blue \times). Both methods overestimate the marginal, but rDMP is closer to the true value (the line $y = x$) for every node.

To investigate how rDMP compares with the pair approximation across a broader range of parameters, in Fig. 5 we vary the ratio between waning rate ρ and the transmission rate λ . Since we can always rescale time by multiplying λ and ρ by the same constant, we do this by holding $\lambda = 0.1$ as before and varying ρ . We then measure the difference in the L_1 error of the two methods, $L_1^{\text{rDMP}} - L_1^{\text{pair}}$.

In the blue region, rDMP is more accurate than the pair approximation; in the red region, it is less so. We see that rDMP is more accurate except at early times (as in Fig. 3) or when ρ is small compared to λ , i.e., if the model is close to the SI model where infectious nodes rarely become susceptible again. The intuition for these findings, as discussed in Sec. II,

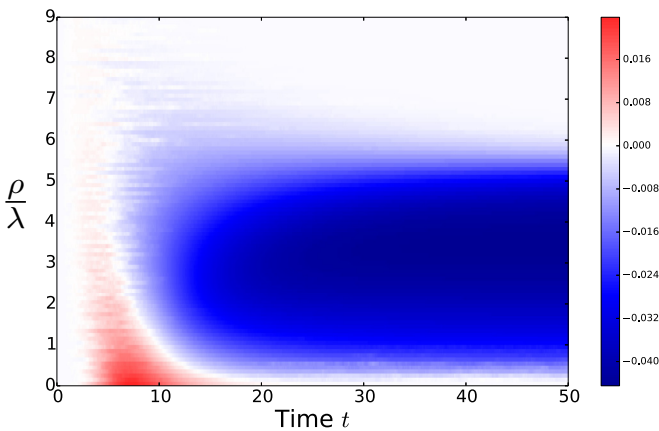


FIG. 5. (Color online) The difference between L_1^{rDMP} and L_1^{pair} on Zachary's karate club for various values of the ratio ρ/λ . We rescale time so that $\lambda = 0.1$ as before. In the blue region, $L_1^{\text{rDMP}} < L_1^{\text{pair}}$ and rDMP is more accurate; in the red region, $L_1^{\text{rDMP}} > L_1^{\text{pair}}$. We see that rDMP is more accurate except at early times or when ρ/λ is small.

is that rDMP assumes that repeated infection events along the same directed edge are independent, and as a result rDMP underperforms pair approximation at early times and when the model is close to SI.

In Fig. 6, we simulate the SIS model on an Erdős-Rényi graph (left figure) with $n = 100$ and average degree 3, with $\lambda = 0.4$, $\rho = 0.1$, and a single initially infectious node. As with the karate club, rDMP does a better job of tracking the true fraction of infectious nodes, except at early times when the pair approximation is superior; in particular, it does a better job of computing the steady-state size of the epidemic.

In Fig. 7 we show results for the SIRS model on Zachary's karate club. As in Fig. 3, on the left we show the marginal probability I_{29} that node 29 is infectious; on the right, we show the L_1 error for I_i averaged over the network. In the insets, we show the marginal probability R_{29} for the recovered state and the corresponding average L_1 error. Here the transmission rate is $\lambda = 0.1$, the waning rate from infectious to recovered is $\rho = 0.05$, and the rate from recovered to susceptible is $\gamma = 0.2$. The initial condition consisted of a single infected node, and Monte Carlo results were averaged over 10^5 runs. As for the SIS model, rDMP is significantly more accurate than the independent node approximation, and is more accurate than the pair approximation except at early times.

We found similar results on many other families of networks, including random regular graphs, random geometric graphs, scale-free networks, Newman-Watts-Strogatz small world networks, and a social network of dolphins [37]. Namely, rDMP outperforms the first-order approximation where nodes are independent, and outperforms the pair approximation across a wide range of parameters and times.

V. LINEAR STABILITY, EPIDEMIC THRESHOLDS, AND RELATED WORK

Systems of differential equations for rDMP, such as Eq. (3), do not appear to have a closed analytic form due to their nonlinearities. On the other hand, we can compute quantities such as epidemic thresholds by linearizing around a stationary point, such as $\{I_{j \rightarrow i}^* = 0\}$, where the initial outbreak is small. Given a perturbation $\epsilon_{j \rightarrow i} = I_{j \rightarrow i} - I_{j \rightarrow i}^*$, the linear stability of the system, i.e., whether or not $\epsilon_{j \rightarrow i}$ diverges in time, is governed by the eigenvalues of the Jacobian matrix \mathbf{J} of the right-hand side of Eq. (3) at the stationary point I_i^* . The Jacobian for Eq. (3) at $\{I_{j \rightarrow i}^*\}$ is

$$\mathbf{J}_{(j \rightarrow i), (k \rightarrow j')} = -\delta_{kj} \delta_{ij'} \rho + \lambda(1 - I_j^*) \mathbf{B}_{(j \rightarrow i), (k \rightarrow j')}, \quad (9)$$

where

$$\mathbf{B}_{(j \rightarrow i), (k \rightarrow j')} = \delta_{jj'}(1 - \delta_{ik}), \quad (10)$$

and δ_{ij} is 1 if $i = j$, and 0 otherwise. This definition of \mathbf{B} is another way of saying that the edge $k \rightarrow j$ influences edges $j \rightarrow i$ for $i \neq k$, but does not backtrack to k . This corresponds to our assumption that infections, for instance, do not bounce from k to j and back again and create an ‘‘echo chamber effect.’’ For this reason, \mathbf{B} is also known in the literature as the nonbacktracking matrix [38] or the Hashimoto matrix [39].

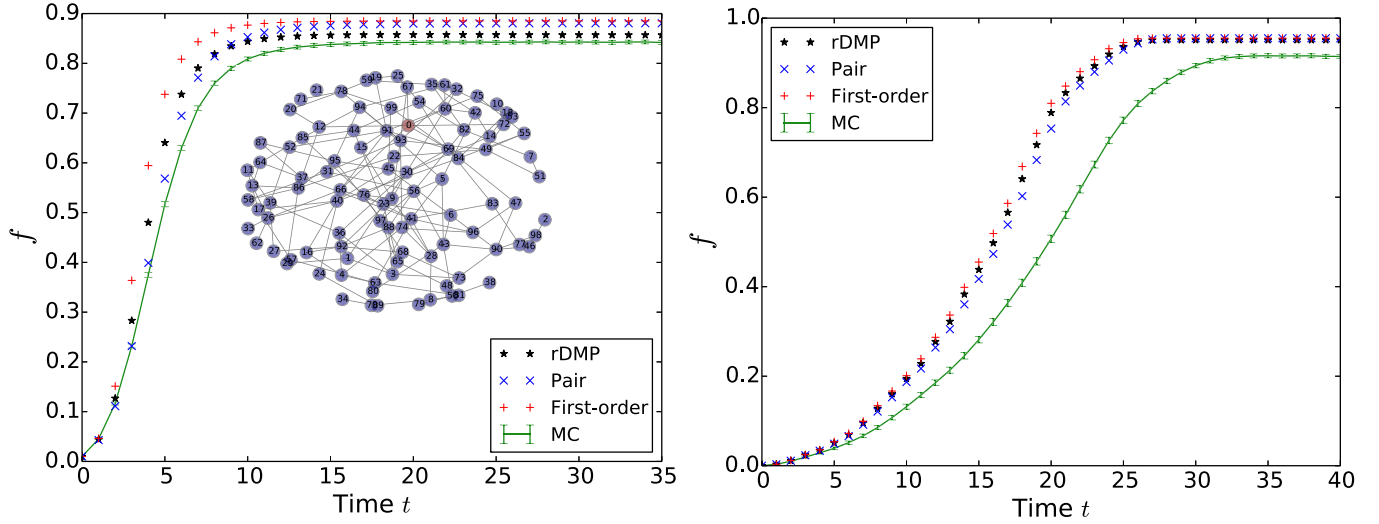


FIG. 6. (Color online) On the left, we show the fraction f of infectious nodes as a function of time in the SIS model on an Erdős-Rényi graph (inset) with $n = 100$ and average degree 3. Here, $\lambda = 0.4$, $\rho = 0.1$, and the initial condition consists of a single infectious node (colored red). Monte Carlo results were averaged over 10^3 independent runs. On the right is the same as the figure on the left, except we show the results from a larger but random geometric graph with 10^5 nodes, where the average clustering coefficient is 0.6. Monte Carlo results here were averaged over 100 independent runs. Except at early times, rDMP tracks the true trajectory more closely.

Now, for a small perturbation $\vec{\epsilon}$ away from a stationary point $\{I_{j \rightarrow i}^*\}$, the linearized system of Eq. (3) becomes

$$\frac{d\vec{\epsilon}}{dt} = \mathbf{J}\vec{\epsilon}. \quad (11)$$

If \mathbf{J} has any eigenvalues with positive real parts, then $\|\vec{\epsilon}(t)\|$ grows exponentially in time. So, the fixed point $\{I_{j \rightarrow i}^*\}$ is stable as long as the leading eigenvalue J_1 of \mathbf{J} has negative real parts.

One trivial, but important, stationary point to test is $I_{j \rightarrow i}^* = 0$ for all edges. A small perturbation around $\vec{0}$ corresponds to a small initial probability that each node is infected. From

Eq. (9), \mathbf{J} becomes

$$\mathbf{J} = \lambda \left(\mathbf{B} - \frac{\rho}{\lambda} \mathbf{1} \right), \quad (12)$$

where $\mathbf{1}$ is the $2m \times 2m$ identity matrix. So, the leading eigenvalue of \mathbf{J} becomes positive when the largest eigenvalue B_1 of \mathbf{B} is greater than ρ/λ . In other words, if

$$R_0 = \frac{\lambda}{\rho} B_1 \geq 1, \quad (13)$$

where R_0 is the reproductive number, even a small initial probability of infection will lead to a widespread endemic

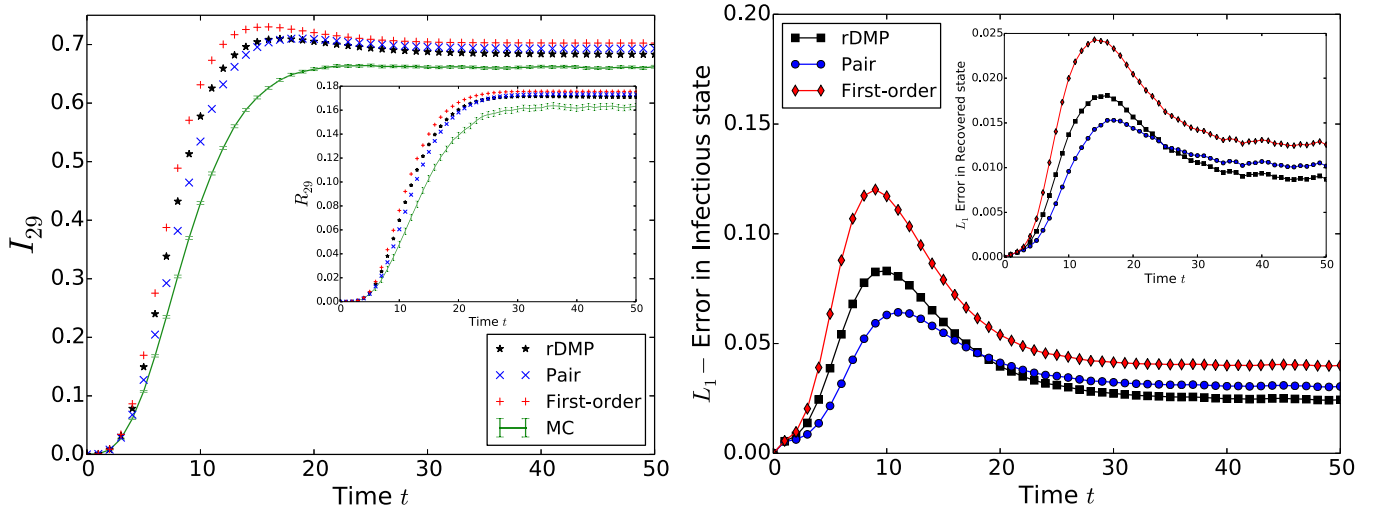


FIG. 7. (Color online) The SIRS model on the karate club. On the left, we show the true and estimated marginal probability that a node 29 is infectious (main figure) or recovered (inset) as a function of time. On the right is the average L_1 error for the infectious and recovered (inset) states. The transmission rate is $\lambda = 0.1$, and the transition rates from infectious to recovered and from recovered to susceptible are $\rho = 0.05$ and $\gamma = 0.2$, respectively. Node 0 (colored red) was initially infected. Monte Carlo results were averaged over 10^5 runs. As for the SIS model, rDMP is significantly more accurate than the first-order model where nodes are independent and is more accurate than the pair approximation except at early times.

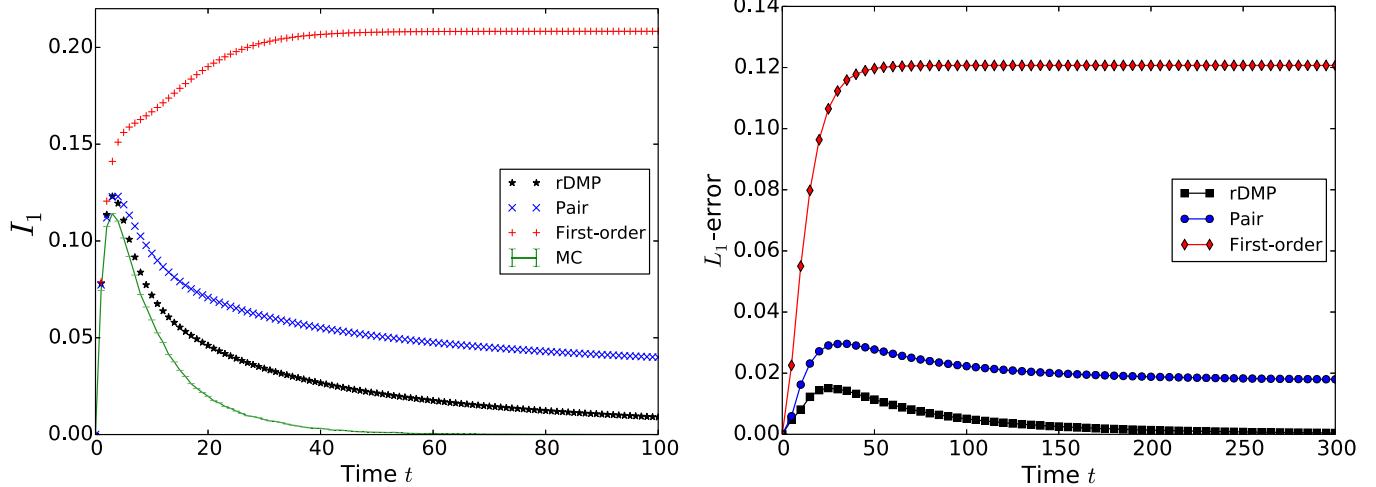


FIG. 8. (Color online) Same as described in the caption of Fig. 3, but with transmission rate $\lambda = 0.1$ and waning rate $\rho = 0.54$. A well-known upper bound on the epidemic threshold of the SIS model can be computed from the leading eigenvalue A_1 of the adjacency matrix (the Jacobian matrix of first-moment-closure approach) of a network. In other words, if $\frac{\rho}{\lambda} < A_1$, it is known from the first-moment-method that an infection-free state becomes unstable and epidemics become widespread and endemic. Here we show the results from SIS model in Zachary’s karate club, where $A_1 \approx 6.7$. Even though $\frac{\rho}{\lambda} = 5.4 < A_1$, which is well below the threshold from the first-moment method, the contagion fades away eventually, which is correctly captured by our DMP approach.

state, where the infection becomes extensive. If Eq. (13) does not hold, a small initial probability of infection will instead decay back to an infection-less state.

Since \mathbf{B} is not symmetric, not all of its eigenvalues are real. However, by the Perron-Frobenius theorem, its leading eigenvalue is real; moreover, it is upper bounded by A_1 , the leading eigenvalue of the adjacency matrix \mathbf{A} . Interestingly, if we examine the linear stability of the first-order approximation where nodes are independent [17], the epidemic threshold for the SIS model is given by

$$\frac{\lambda}{\rho} A_1 \geq 1. \quad (14)$$

Since $B_1 \leq A_1$, the threshold Eq. (13) gives a better upper bound for the true epidemic threshold than we would get from the first-order approximation. In Fig. 8, we show the results of SIS dynamics in Zachary’s karate club closer to epidemic transition, where first-order approximation is highly error-prone.

A similar threshold for the SIR model in sparse networks, or equivalently for percolation, using B_1 was recently demonstrated in Ref. [40]. (We note that when backtracking is allowed, it has important consequences for epidemic thresholds on power-law networks [41].)

Whereas the leading eigenvector of \mathbf{B} governs the epidemic threshold, the spectral gap between \mathbf{B} ’s top two eigenvectors governs how quickly the epidemic converges to the leading behavior (at least until we leave the linear regime). Qualitatively, this depends on bottlenecks in the network such as those due to community structure, where an epidemic spreads quickly in one community but then takes a longer time to cross over into another. Indeed, the second eigenvector of the nonbacktracking matrix \mathbf{B} was recently used to detect community structure [38].

Similarly, just as the leading eigenvector of \mathbf{B} was recently shown to be a good measure of importance or “centrality” of a node [42], it may be helpful in identifying “superspreaders”—

nodes where an initial infection will generate the largest outbreak—and be the most likely to lead to a widespread epidemic.

VI. CONCLUSION

Modern epidemiological studies often require recurrent models, where nodes can return to their previous inhabited states multiple times. For example, consider diseases such as influenza where individuals are infected multiple times throughout their lives, or whooping cough where vaccine effectiveness wanes over time; in both cases, individuals return to the susceptible class. In this paper we have extended DMP to recurrent epidemic models. Our rDMP approach defines messages on the directed edges of a network in such a way as to prevent signals, such as the spread of infection, from backtracking immediately to the node that they came from. By preventing these “echo chamber effects,” rDMP obtains good estimates of the time-varying marginal probabilities on a wide variety of networks, estimating both the fraction of infectious individuals in the entire network, and the probabilities that individual nodes become infected.

Like the pair approximation, rDMP takes correlations between neighboring nodes into account. However, our experiments show that rDMP is more accurate than the pair approximation for a wide variety of network structures and parameters. Moreover, rDMP is computationally less expensive than the pair approximation, especially for complex epidemic models with a large number of states, using $O(mk)$ instead of $O(mk^2)$ variables for models with k states on networks with m edges.

Finally, rDMP is conceptually simple, allowing the user to immediately write down the system of differential equations for a wide variety of epidemic models, such as those with multiple stages of infection or immunity [43,44], or those with multiple interacting diseases [45,46]. We expect that given its

simplicity and accuracy, it will be an attractive option for future epidemiological studies.

ACKNOWLEDGMENTS

This work is supported by AFOSR and DARPA under Grant No. FA9550-12-1-0432. M.S. acknowledges support

through a graduate fellowship at the Santa Fe Institute. S.V.S. was supported by the Santa Fe Institute and the Omidyar Group. C.M. is supported in part by the John Templeton Foundation. We are grateful to Mason Porter and Joel Miller for helpful conversations regarding recurrent state epidemic models.

-
- [1] N. T. J. Bailey, *The Mathematical Theory of Infectious Diseases and Its Applications* (Hafner Press, New York, 1975).
- [2] R. M. Anderson and R. M. May, *Infectious Diseases of Humans* (Oxford University Press, Oxford, 1991).
- [3] M. Granovetter, Threshold models of collective behavior, *Am. J. Sociol.* **83**, 1420 (1978).
- [4] M. Granovetter, The strength of weak ties, *Am. J. Sociol.* **78**, 1360 (1973).
- [5] J. H. Miller and S. E. Page, The standing ovation problem, *Complexity* **9**, 8 (2004).
- [6] B. Gonçalves, N. Perra, and A. Vespignani, Modeling users' activity on Twitter networks: Validation of Dunbar's number, *PLoS ONE* **6**, e22656 (2011).
- [7] R. M. May and A. G. Haldane, Systemic risk in banking ecosystems, *Nature* **469**, 351 (2011).
- [8] F. Caccioli, M. Shrestha, C. Moore, and J. D. Farmer, Stability analysis of financial contagion due to overlapping portfolios, *J. Bank. Finance* **46**, 233 (2014).
- [9] P. Bak, K. Chen, and C. Tang, A forest-fire model and some thoughts on turbulence, *Phys. Lett. A* **147**, 297 (1990).
- [10] B. Drossel and F. Schwabl, Self-organized critical forest-fire model, *Phys. Rev. Lett.* **69**, 1629 (1992).
- [11] P. Grassberger, Critical behavior of the Drossel-Schwabl forest fire model, *New J. Phys.* **4**, 17 (2002).
- [12] M. Mézard and A. Montanari, *Information, Physics, and Computation* (Oxford University Press, Oxford, 2009).
- [13] C. Moore and S. Mertens, *The Nature of Computation* (Oxford University Press, Oxford, 2011).
- [14] R. Morris, Zero-temperature Glauber dynamics on Z^d , *Prob. Theory Rel. Fields* **149**, 417 (2009).
- [15] R. I. M. Dunbar, Neocortex size as a constraint on group size in primates, *J. Human Evol.* **22**, 469 (1992).
- [16] L. A. Meyers, Contact network epidemiology: Bond percolation applied to infectious disease prediction and control, *Bull. Am. Math. Soc.* **44**, 63 (2007).
- [17] M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, Oxford, 2010).
- [18] A. Y. Lokhov, M. Mézard, H. Ohta, and L. Zdeborová, Inferring the origin of an epidemic with dynamic message-passing algorithm, *Phys. Rev. E* **90**, 012801 (2014).
- [19] F. Altarelli, A. Braunstein, L. Dall'Asta, A. Ingrosso, and R. Zecchina, The zero-patient problem with noisy observations, *J. Stat. Mech.* (2014) P10016.
- [20] F. Altarelli, A. Braunstein, L. Dall'Asta, J. R. Wakeling, and R. Zecchina, Containing Epidemic Outbreaks by Message-Passing Techniques, *Phys. Rev. X* **4**, 021024 (2014).
- [21] F. Altarelli, A. Braunstein, L. Dall'Asta, and R. Zecchina, Optimizing spread dynamics on graphs by message passing, *J. Stat. Mech.* (2013) P09011.
- [22] F. Altarelli, A. Braunstein, A. Ramezanzpour, and R. Zecchina, Stochastic optimization by message passing, *J. Stat. Mech.* (2011) P11009.
- [23] M. A. Porter and J. P. Gleeson, Dynamical systems on networks: A tutorial. [arXiv:1403.7663](https://arxiv.org/abs/1403.7663) (2014).
- [24] J. Pearl, Reverend Bayes on inference engines: A distributed hierarchical approach, *AAAI Proc.* **82**, 133–136 (1982).
- [25] 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).
- [26] B. Karrer and M. E. J. Newman, Message passing approach for general epidemic models, *Phys. Rev. E* **82**, 016101 (2010).
- [27] M. Shrestha and C. Moore, Message passing approach for threshold models of behavior in networks, *Phys. Rev. E* **89**, 022805 (2014).
- [28] F. Altarelli, A. Braunstein, L. Dall'Asta, and R. Zecchina, Large deviations of cascade processes on graphs, *Phys. Rev. E* **87**, 062115 (2013).
- [29] A. Y. Lokhov, M. Mézard, and L. Zdeborová, Dynamic message-passing equations for models with unidirectional dynamics, *Phys. Rev. E* **91**, 012811 (2015).
- [30] D. J. D. Earn, J. Dushoff, and S. A. Levin, *Ecology and evolution of the flu*, *Trends Ecol. Evol.* **17**, 334 (2002).
- [31] M. G. M. Gomes, L. J. White, G. F. Medley, Infection, reinfection, and vaccination under suboptimal immune protection: Epidemiological perspectives, *J. Theor. Biol.* **228**, 539 (2004).
- [32] M. R. Golden, W. L. H. Whittington, H. H. Handsfield, J. P. Hughes, W. E. Stamm, M. Hogben, A. Clark, C. Malinski, J. R. L. Hesters, K. K. Thomas, and K. K. Holmes, Effect of expedited treatment of sex partners on recurrent or persistent gonorrhea or chlamydial infection, *New England J. Med.* **352**, 676 (2005).
- [33] G. M. Jeffery, Epidemiological significance of repeated infections with homologous and heterologous strains and species of Plasmodium, *J. Bull. World Health Org.* **35**, 873 (1966).
- [34] P. H. Conway, A. Cnaan, T. Zaoutis, and B. V. Henry, R. W. Grundmeier, and R. Keren, Recurrent urinary tract infections in children: Risk factors and association with prophylactic antimicrobials, *J. Am. Med. Assoc.* **298**, 179 (2007).
- [35] L. M. Drusin, B. G. Ross, K. H. Rhodes, A. N. Krauss, and R. A. Scott, Nosocomial ringworm in a neonatal intensive care unit: A nurse and her cat, *Infect. Control* **21**, 605 (2000).
- [36] W. W. Zachary, An information flow model for conflict and fission in small groups, *J. Anthropol. Res.* **33**, 452 (1977).
- [37] D. Lusseau, K. Schneider, O. J. Boisseau, P. Haase, E. Sloaten, and S. M. Dawson, The bottlenecked dolphin community of Doubtful Sound features a large proportion of long-lasting associations, *Behav. Ecol. Sociobiol.* **54**, 396 (2003).

- [38] F. Krzakala, C. Moore, E. Mossel, J. Neeman, A. Sly, L. Zdeborová, and P. Zhang, Spectral redemption in clustering sparse networks, *Proc. Natl. Acad. Sci. USA* **110**, 20935 (2013).
- [39] K. Hashimoto, Zeta functions of finite graphs and representations of p -adic groups, *Adv. Studies Pure Math.* **15**, 211 (1989).
- [40] B. Karrer, M. E. J. Newman, and L. Zdeborová, Percolation on Sparse Networks, *Phys. Rev. Lett.* **113**, 208702 (2014).
- [41] S. Chatterjee and R. Durrett, Contact processes on random graphs with power law degree distributions have critical value 0, *Ann. Prob.* **37**, 2332 (2009).
- [42] T. Martin, X. Zhang, and M. E. J. Newman, Localization and centrality in networks, *Phys. Rev. E* **90**, 052808 (2014).
- [43] S. Melnik, J. A. Ward, J. P. Gleeson, and M. A. Porter, Multi-stage complex contagions, *Chaos* **23**, 013124 (2013).
- [44] J. C. Miller and E. M. Volz, Incorporating disease and population structure into models of SIR disease in contact networks, *PLoS ONE* **8**, e69162 (2013).
- [45] B. Karrer and M. E. J. Newman, Competing epidemics on complex networks, *Phys. Rev. E* **84**, 036106 (2011).
- [46] J. C. Miller, Cocirculation of infectious diseases on networks, *Phys. Rev. E* **87**, 060801 (2013).