Interacting epidemics on overlay networks

Sebastian Funk and Vincent A. A. Jansen
School of Biological Sciences, Royal Holloway, University of London, Egham, Surrey TW20 0EX, United Kingdom
(Received 29 October 2009; published 31 March 2010)

The interaction between multiple pathogens spreading on networks connecting a given set of nodes presents an ongoing theoretical challenge. Here, we aim to understand such interactions by studying bond percolation of two different processes on overlay networks of arbitrary joint degree distribution. We find that an outbreak of a first pathogen providing immunity to another one spreading subsequently on a second network connecting the same set of nodes does so most effectively if the degrees on the two networks are positively correlated. In that case, the protection is stronger the more heterogeneous the degree distributions of the two networks are. If, on the other hand, the degrees are uncorrelated or negatively correlated, increasing heterogeneity reduces the potential of the first process to prevent the second one from reaching epidemic proportions. We generalize these results to cases where the edges of the two networks overlap to arbitrary amount, or where the immunity granted is only partial. If both processes grant immunity to each other, we find a wide range of possible situations of coexistence or mutual exclusion, depending on the joint degree distribution of the underlying networks and the amount of immunity granted mutually. These results generalize the concept of a coexistence threshold and illustrate the impact of large-scale network structure on the interaction between multiple spreading agents.

DOI: 10.1103/PhysRevE.81.036118 PACS number(s): 89.75.Hc, 87.23.Ge, 05.70.Fh, 64.60.aq

I. INTRODUCTION

The view of a population as a set of individuals connected by a social network has become the paradigm for studying the spread of contagious processes in structured populations. Work in this area has led to the development of a number of techniques yielding insights on epidemic thresholds in such networks [1–3].

In many cases, two or more processes interact on two, not necessarily equal networks connecting the same set of nodes. Two different diseases using potentially different routes of spread can interact through suppression of the immune system by one to facilitate infection with the other or by one granting partial immunity to the other [4–7]. Similar processes can play a role for two competing strains of the same disease [8,9] and the implications of such interactions present an ongoing theoretical challenge.

The interest in overlay networks extends beyond the interaction between only pathogens in the biological sense. Recently, the interaction between epidemic networks of disease spread and social networks of the spread of awareness [10,11] or influence [12] has attracted greater interest. In a computing framework, the interaction between mobile phone viruses exploiting different routes of spread has been recognized as a threat [13]. On the other hand, methods to mitigate the spread of viruses using counterviruses have been suggested [14], in which case a carefully chosen network could maximize the efficiency of such a strategy.

So far, studies of interacting pathogens have largely ignored the impact of network structure. In a notable exception, a study has demonstrated the existence of a *coexistence threshold* for the same pathogen being able to form two epidemics on the same network if those having experienced the first infection are completely immune to the second one [15]. For this to happen, the transmission probability must be high enough for the process to be able to form an epidemic at all (i.e., it must be greater than the epidemic threshold), but also low enough for the first wave to be able to leave a large enough fraction of the network susceptible to the disease (i.e., smaller than the coexistence threshold).

In this contribution, we generalize the idea of a coexistence threshold presented in [15] to allow us to determine conditions for coexistence of two processes of different transmissibility interacting on two networks of arbitrary overlap. We study the effect of correlations between the two networks of varying heterogeneity, as well as considering situations where the two processes grant only partial immunity or, conversely, facilitate infection with the other.

II. METHODS

In the stochastic susceptible-infected-recovered (SIR) model on a network, an infected node can infect connected susceptible nodes with an infection rate $\hat{\beta}$. Once a node is infected, it remains so for an exponentially distributed period of average γ^{-1}, after which it recovers, and can neither be infected nor infect others anymore. If the infection events are taken to happen as independent Poisson processes of intensity $\hat{\beta}$, this is equivalent to saying that each possible infection between two connected nodes is realized independently with probability $T=\hat{\beta}/(\hat{\beta}+\gamma)$. This consideration includes the SI (susceptible-infected) model, which does not allow for recovery of infected nodes, as a special case of setting $\gamma=0$, so that $T=1$.

The distribution of outbreak sizes of the stochastic SIR model on random networks of a given degree distribution can be calculated by exploiting an isomorphism with the distribution of connected component sizes in bond percolation theory [1]. If edges on a random network of infinite size are selected or "occupied" independently with probability T, the resulting connected components of occupied edges (i.e.,

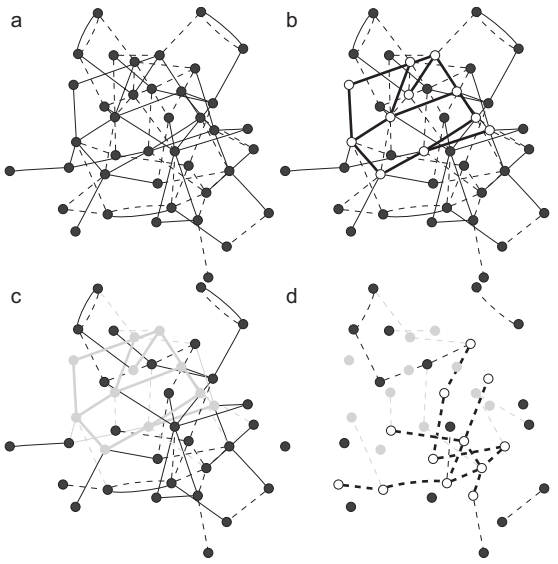

FIG. 1. The method we use to study interacting epidemics on overlay networks. (a) We start with a set of nodes connected by two random networks Γ_1 (solid lines) and Γ_2 (dashed lines). (b) We then select or "occupy" edges on Γ_1 independently with probability T_1. If T_1 is large enough, a giant connected component of occupied edges emerges (thick solid lines), representing a large outbreak and occupying a fraction $S_1(T_1)$ of the nodes (white dots). (c) The nodes in that component (gray dots) are considered immune and removed from Γ_2, leaving the residual network of Γ_1 on Γ_2 (black dashed lines). (d) Edges on that network are then occupied with probability T_2, with a giant connected component of occupied edges (thick dashed lines) emerging if T_2 is larger than a threshold which depends on T_1 and the structure of Γ_1 and Γ_2. If the immunity granted by the first epidemic is only partial the gray nodes are not removed in step c, but edges connecting to them on Γ_2 (gray dashed lines) are assumed to be occupied with a reduced probability.

subnetworks in which each node can be reached from each other one by following only edges which have been selected) have the same size distribution as outbreaks of the stochastic SIR model on a random network of the same degree distribution.

This method can be extended to interacting epidemics on overlay networks where one epidemic grants immunity to the other and the two spread on different networks connecting the same nodes (see Fig. 1 for illustration). Consider two infinitely sized networks Γ_1 and Γ_2 which connect the same sets of vertices. We first occupy edges on Γ_1 independently with probability T_1. If T_1 is greater than a critical probability $T_{C,1}$, a *giant connected occupied component* (GCOC) emerges on Γ_1, occupying a fraction $S_1(T_1)$ of all nodes. This is equivalent to an SIR epidemic on a random network with the same degree distribution as Γ_1, causing a large outbreak with probability $S_1(T_1)$ which, if it happens, infects a fraction $S_1(T_1)$ of the network [1]. We then remove all the nodes in the GCOC of Γ_1 from the network and occupy the remaining edges on Γ_2 with probability T_2. In other words, we concentrate on the residual network of Γ_1 on Γ_2, that is the network of nodes which are not part of the GCOC on Γ_1, and which are connected on Γ_2. This is equivalent to the first SIR epidemic providing immunity to a second SIR process which starts after the first has run its course, now spreading on a network represented by Γ_2, connecting the same nodes but possibly forming different paths. We can then study the emergence of a GCOC on Γ_2 and calculate the corresponding probability of a large outbreak as a function of T_1 and the structure of the two networks. We will begin with the simplest case of two networks with a joint degree distribution which do not overlap, but later will extend this to more general cases.

Note that we here only consider the impact of large outbreaks of one process on the other, where the second one starts at a random location *not* affected by the first. This is because in the limit of networks of infinite size the size of small outbreaks, and therefore the probability for the two epidemics to meet in nodes not in the GCOC, tends to zero. There can, however be a local interaction if the two are assumed to start at the same location, and if there is overlap between the two networks [see [10]]. In that case, one needs a dynamic approach and the static approximation taken here is of little use.

Let us consider the general case where the degrees k_1 and k_2 of the nodes which are connected by the two graphs are distributed with a joint degree distribution $p(k_1,k_2)$ as generated by the joint (J) generating function

$$G_0^{\mathrm{J}}(x,y) = \sum_{k_1} \sum_{k_2} p(k_1,k_2) x^{k_1} y^{k_2}, \tag{1}$$

and marginal degree distributions $p_1(k_1)$ and $p_2(k_2)$ generated by

$$G_{0,1}(x) = \sum_{k_1} p_1(k_1) x^{k_1} = \sum_{k_1} \sum_{k_2} p(k_1,k_2) x^{k_1},$$

$$G_{0,2}(x) = \sum_{k_2} p_2(k_2) x^{k_2} = \sum_{k_1} \sum_{k_2} p(k_1,k_2) x^{k_2}. \tag{2}$$

Furthermore, we for now assume that the two networks do not have any common edges and that the processes mutually exclude each other, i.e., any node can be infected with only one of the two processes and once infected with one is immune to the other.

If $T_1 > T_{C,1}$, with $T_{C,1}$ given by

$$T_{C,1} = \frac{\langle k_1 \rangle}{\langle k_1^2 \rangle - \langle k_1 \rangle}, \tag{3}$$

or, analogously, $R_0 > 1$, with the basic reproductive number R_0 of the network given by [16–19]

$$R_0 = T_{C,1}\left(\langle k_1 \rangle - 1 + \frac{\mathrm{Var}(k_1)}{\langle k_1 \rangle} \right), \tag{4}$$

the GCOC of Γ_1 occupies a fraction $S_1(T_1)$ [1]

$$S_1(T_1) = 1 - G_{0,1}(u_1;T_1), \tag{5}$$

where u_1 is the solution of

$$u_1 = G_{1,1}(1 - T_1 + T_1 u_1), \tag{6}$$

$G_{1,1}(x)$ is the generating function of the excess degree of a node at the end of a randomly selected edge on Γ_1,

$$G_{1,1}(x) = \frac{1}{\langle k_1 \rangle} \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} (k_1+1) p(k_1+1, k_2) x^{k_1}. \quad (7)$$

and $G_{0,1}(x;T_1)$ is the probability generating function (pgf) for a random node to have n of its edges occupied [1],

$$G_{0,1}(x;T) = G_{0,1}[1+(x-1)T]. \quad (8)$$

To distinguish the generating functions $G_{\cdot,\cdot}(x;T)$ which concern occupied edges from those $G_{\cdot,\cdot}(x)$ which concern all edges, with the two connected to each other by relations such as Eq. (8), note that the T is separated by a semicolon, a convention which we will follow throughout this manuscript.

Denoting with $\epsilon_{k_1}(T_1)$ the probability of a randomly chosen node of degrees k_1 and k_2 to be part of the GCOC on Γ_1, we have [1]

$$\mathrm{P}(\text{not in gcoc of } \Gamma_1 | k_1) = 1 - \epsilon_{k_1}(T_1) = (1 - T_1 + T_1 u_1)^{k_1}. \quad (9)$$

At the same time, the probability for m of the k_2 neighbors on Γ_2 of a randomly selected node not in the GCOC of Γ_1 not to be part of the GCOC on Γ_1 either is

$$\mathrm{P}(\text{connected to } m \text{ not in gcoc of } \Gamma_1 | k_2) = \binom{k_2}{m} w_1^m (1-w_1)^{k_2-m}, \quad (10)$$

with w_1 given by

$$w_1 = \frac{1}{\langle k_2 \rangle} \sum_{k_1,k_2} p(k_1,k_2) k_2 (1 - T_1 + T_1 u_1)^{k_1} = \frac{1}{\langle k_2 \rangle} \frac{\partial}{\partial y} G_0^{\mathrm{J}}(1 - T_1 + T_1 u_1, 1), \quad (11)$$

where $\langle\ \rangle$ denotes the average on the network, understood to be operating on the joint degree distribution $p(k_1,k_2)$

$$\langle f(k_1,k_2) \rangle = \sum_{k_1,k_2} p(k_1,k_2) f(k_1,k_2). \quad (12)$$

w_1 is the probability of a node arrived at following a random edge on Γ_2 to be part of the GCOC on Γ_1, and it takes into account the fact that the probability of finding a node of degree k_2 following a randomly selected edge on the second network is proportional to $k_2 p(k_1,k_2)$.

The probability of a node of degree k_1 on Γ_1 and k_2 on Γ_2 not to be part of the GCOC on Γ_1 and to be connected to m nodes that are not part of that GCOC either is obtained by multiplying Eqs. (9) and (10), and it is generated by

$$g_0^r(x;k_1,k_2) = \sum_r (1 - T_1 + u_1 T_1)^{k_1} \binom{k_2}{m} w_1^m (1-w_1)^{k_2-m} x^m = (1 - T_1 + u_1 T_1)^{k_1} (1 - w_1 + w_1 x)^{k_2}. \quad (13)$$

The generating function $G_{0,2}^r(x)$ of the residual degree distribution of Γ_1 on Γ_2 is then found by averaging Eq. (13) over all probabilities $p(k_1,k_2)$ which, normalized to $G_{0,2}^r(1)=1$, gives

$$\begin{aligned} G_{0,2}^r(x) &= \frac{1}{G_{0,1}(u_1;T_1)} \sum_m \sum_{k_1} \sum_{k_2} p(k_1,k_2) (1 - T_1 + T_1 u_1)^{k_1} \binom{k_2}{m} \\ &\quad \times w_1^m (1-w_1)^{k_2-m} x^m \\ &= \frac{1}{G_{0,1}(u_1;T_1)} \sum_{k_1}^{\infty} \sum_{k_2}^{\infty} p(k_1,k_2) \\ &\quad \times (1 - T_1 + T_1 u_1)^{k_1} (1 - w_1 + w_1 x)^{k_2} \\ &= \frac{G_0^{\mathrm{J}}(1 - T_1 + T_1 u_1, 1 - w_1 + w_1 x)}{G_0^{\mathrm{J}}(1 - T_1 + T_1 u_1, 1)}. \end{aligned} \quad (14)$$

Analogously to [20], it can be shown that the generating function for the number of edges occupied on Γ_2 connected to a random vertex on the residual network of Γ_1 is

$$G_{0,2}^r(x;T_2) = G_0^r(1 - T_2 + T_2 x), \quad (15)$$

The equations for $G_1^r(x)$ and $G_{1,2}^r(x;T_2)$ for the generating functions of the excess degree distribution and the number of occupied edges connected to a vertex at the end of a random edge on the residual network follow accordingly by weighing the summands in Eqs. (14) and (15) with k_2.

III. RESULTS

Using the methods developed in the last section, we now derive results for different scenarios of network overlap and immunity.

A. Nonoverlapping networks

Equipped with all the relevant generating functions, we can proceed just as on a single network to calculate epidemiologically relevant quantities. Since part of the network is shielded from the epidemic on Γ_2 because it is already part of the GCOC on Γ_1, the critical transmission probability $T_{\mathrm{C},2}$ of the second network will be higher than on the first one, which led [15] to call it a second, or *coexistence* threshold in the special case where two processes spread on the same network, and with the same transmission probability T. If the transmission probability T_2 is greater than $T_{\mathrm{C},2}$ a GCOC will emerge on the second network and the two processes can coexist. As only a fraction $S_1(T_1)=1-G_{0,1}(u_1;T_1)$ of all nodes in the network are part of the residual network of Γ_1 on Γ_2, the size of the GCOC $S_2(T_2)$ as determined by applying Eq. (5) to Γ_2 using Eq. (15), needs to be multiplied with this fraction to give the fraction $S_2(T_2)$ of all nodes being part of the GCOC in the second network

$$S_2(T_2) = G_{0,1}(u_1;T_1)[1 - G_{0,2}^r(u_2;T_2)]. \quad (16)$$

The basic reproductive number of the second process spreading on Γ_2 is

$$R_{0,2} = \left. \frac{\partial G_{1,2}^r(x;T_2)}{\partial x} \right|_{x=1} = T_2 G_1^{r\prime}(1) = T_2 \frac{\langle k_2 (k_2-1)(1 - T_1 + T_1 u_1)^{k_1} \rangle}{\langle k_2 \rangle}. \quad (17)$$

The threshold for the second process to have a GCOC is the transmission probability at which $R_{0,2}=1$

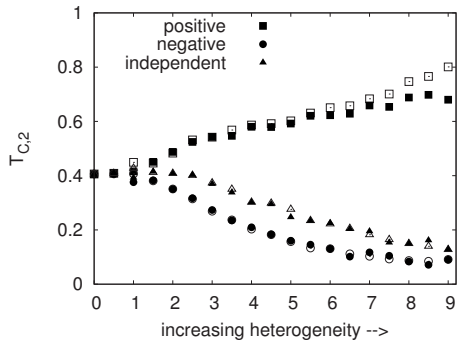

FIG. 2. The critical transmission probability $T_{C,2}$ as a function of heterogeneity for $T_{C,1}=0.25$, expressed as the standard deviation in the degree distribution for an average of $\langle k\rangle=6$, for positive (squares), negative (circles), and no (triangles) correlation between degrees of nodes on both networks. The critical transmission as determined from simulations of the dynamic SIR model on a finite network of $N=10\,000$ nodes is shown for comparison (open symbols). Where only a filled dot can be seen, the open one is in the same position. Details on how the networks were generated, and how the critical transmission probabilities estimated from simulations, can be found in the appendix.

$$T_{C,2}=\frac{\langle k_2\rangle}{\langle k_2(k_2-1)(1-T_1+T_1u_1)^{k_1}\rangle}. \tag{18}$$

If the degrees on the two networks are independently distributed with $p_1(k_1)$ and $p_2(k_2)$, we have that $p(k_1,k_2)=p_1(k_1)p_2(k_2)$ and can rewrite Eq. (14) to

$$\begin{aligned}G_0^r(x)&=\frac{1}{G_{0,1}(u_1;T_1)}\sum_m^\infty\sum_{k_1}^\infty\sum_{k_2}^\infty p_1(k_1)p_2(k_2)\binom{k_2}{m}\\&\quad\times w_1^m(1-w_1)^{k_2-m}(1-T_1+T_1u_1)^{k_1}\\&=\frac{G_{0,1}(1-T_1+T_1u_1)G_{0,2}(1-w_1+w_1x)}{G_{0,1}(1-T_1+T_1u_1)}\\&=G_{0,2}(1-w_1+w_1x),\end{aligned} \tag{19}$$

with the other generating functions again following analogously. The basic reproductive number of the second process in this case becomes

$$R_{0,2}=\left.\frac{\partial G_{1,2}^r(x;T_2)}{\partial x}\right|_{x=1}=G_{0,1}(u_1;T_1)T_2\frac{\langle k_2(k_2-1)\rangle}{\langle k_2\rangle}. \tag{20}$$

Equation (20) is the standard expression for the basic reproductive number on a single network as given by Eq. (4), multiplied by $G_{0,1}(u_1;T_1)$, the probability of a randomly chosen node not to be part of the GCOC on the first network. Analogously, the critical transmission probability $T_{C,2}$ is

$$T_{C,2}=\frac{1}{G_{0,1}(u_1;T_1)}\frac{\langle k_2\rangle}{\langle k_2(k_2-1)\rangle}. \tag{21}$$

We can use the above relations to study the effect of different types of correlation between degrees on both networks. Figure 2 illustrates the impact of a first spreading process on the critical transmission probability $T_{C,2}$ of a second which the first one provides immunity to, for networks of a varying amount of heterogeneity in their degree distribution, generated using a method outlined in the Appendix. If the degrees are distributed independently, $T_{C,2}$ decreases as heterogeneity increases, making it easier for the second process to invade the population with the first one already established. This effect is amplified if the correlation between degrees is negative, i.e., if nodes with a high degree on Γ_1 tend to have small degree on Γ_2 and vice versa. This is because the nodes which have a high degree on Γ_1, and therefore have a high probability of being in the GCOC, are not hubs on Γ_2 and therefore not so relevant for the formation of a GCOC on Γ_2. If, on the other hand, the correlation is positive, i.e., the nodes with high degree on one network tend to have high degree on the other, it becomes harder for the second process to invade as heterogeneity increases because the second process finds its hubs already occupied by the first. Figure 2 also shows comparisons of the results obtained using our percolation-based method with critical transmission probabilities estimated from stochastic simulations on finite networks, with particularly good agreement on networks of relatively low heterogeneity.

B. Networks of arbitrary overlap

If the networks are not completely disjunct, in the sense that some vertices are connected by an edge on both networks, the degree distribution of the residual network $G_{0,2}^r$ changes. For a node connected to k_b other nodes on both Γ_1 and Γ_2, the probability of being connected to n of them on the residual network of Γ_1 on Γ_2 is [15]

$$\binom{k_b}{n}(u_1)^n[(1-T_1)(1-u_1)]^{k_b-n}, \tag{22}$$

which is the probability u_1 of each of the m nodes arrived at following an edge on Γ_1 not to be part of the GCOC times the probability $(1-u_1)(1-T_1)$ of each of the remaining nodes to be part of the outbreak, $(1-u_1)$, but without the edge being occupied, $(1-T_1)$. The probabilities of Eq. (22) are generated by

$$\begin{aligned}g_{0,2}^{r,b}(x;k_b)&=\sum_n\binom{k_b}{n}(u_1)^m[(1-T_1)(1-u_1)]^{k_b-n}x^n\\&=\sum_n\binom{k_b}{n}(xu_1)^n[(1-T_1)(1-u_1)]^{k_b-n}\\&=[(1-T_1)(1-u_1)+xu_1]^{k_b}.\end{aligned} \tag{23}$$

Remembering that the generating function of the sum of two independently distributed random variables is the product of the corresponding generating functions [21], we obtain the generating function for being connected to m nodes of the residual network of Γ_1 via edges of Γ_2 only and n nodes on that same network via edges shared by both networks by multiplying $g_{0,2}^r(x;k_1-k_b,k_2-k_b)$ of Eq. (13) with $g_{0,2}^{r,b}(x;k_b)$ of Eq. (23),

$$g_{0,2}^{\text{overlap}}(x,y;k_1,k_2,k_b) = (1-T_1+T_1u_1)^{k_1-k_b}(1-w_1+xw_1)^{k_2-k_b} \times [(1-T_1)(1-u_1)+u_1y]^{k_b}, \quad (24)$$

and the generating function for being connected to a sum $m+n$ via either is obtained by setting $x=y$,

$$g_{0,2}^{\text{overlap}}(x;k_1,k_2,k_b) = g_{0,2}^{\text{overlap}}(x,x;k_1,k_2,k_b) = (1-T_1+T_1u_1)^{k_1-k_b}(1-w_1+x_1w_1)^{k_2-k_b} \times [(1-T_1)(1-u_1)+u_1x]^{k_b}. \quad (25)$$

Now, we could obtain the generating function for a joint degree distribution $p(k_1,k_2,k_b)$ by averaging over k_1, k_2, and k_b. Instead, we here choose to make an approximation to quantify the amount of overlap between the two networks. Let us assume that each pair of vertices connected by an edge on Γ_2 has independent probability $q_{1|2}$ of being also connected on Γ_1. Likewise, each pair of vertices connected by an edge on Γ_1 has independent probability $q_{2|1}$ of being connected on Γ_2 as well. For consistency reasons, we then have

$$q_{1|2}\langle k_2\rangle = q_{2|1}\langle k_1\rangle. \quad (26)$$

Now, for each of the k_2 neighbors of a randomly selected node, there is a probability $q_{1|2}$ that it is also connected to that node on Γ_1. The probability of having k_b of those edges shared with Γ_1 is therefore

$$\binom{k_2}{k_b} q_{1|2}^{k_b}(1-q_{1|2})^{k_2-k_b}. \quad (27)$$

Using Eqs. (25) and (27), we can then obtain the generating function of the residual degree distribution of Γ_1 on Γ_2 for partially overlapping networks by averaging over $p(k_1,k_2)$ and summing over all possible k_b,

$$G_{0,2}^r(x) = \sum_{k_1,k_2} p(k_1,k_2) \sum_{k_b=0}^{k_2} \binom{k_2}{k_b} q_{1|2}^{k_b}(1-q_{1|2})^{k_2-k_b}(1-T_1+T_1u_1)^{k_1-k_b}(1-w_1+x_1w_1)^{k_2-k_b}[(1-T_1)(1-u_1)+u_1x]^{k_b}$$
$$= \sum_{k_1,k_2} p(k_1,k_2)(1-T_1+T_1u_1)^{k_1}\left\{q_{1|2}\frac{(1-T_1)(1-u_1)+u_1x}{1-T_1+T_1u_1}+(1-q_{1|2})(1-w_1+w_1x)\right\}^{k_2}. \quad (28)$$

Note that in this approximation, $p(k_1,k_2)$ is constrained by the value of $q_{1|2}$, for example if $q_{1|2}=1$ we must have $p(k_1,k_2>k_1)=0$ because all nodes connected by a node on Γ_2 are connected on Γ_1, too.

In the limit of $q_{1|2}=0$ (no overlap), we recover our previous result of Eq. (14), whereas if $q_{1|2}=1$ (complete overlap), we have

$$G_{0,2}^r(x) = \sum_{k_1,k_2} p(k_1,k_2)(1-T_1+T_1u_1)^{k_1/k_2}[(1-T_1)(1-u_1)+u_1x]^{k_2}, \quad (29)$$

which, if $p(k_1,k_2)=p(k_1)\delta_{k_1k_2}$, i.e., if the networks are completely equal, recovers the result presented in [15].

The basic reproductive number for partly overlapping networks is

$$R_{0,2}^{\text{overlap}} = \left.\frac{\partial G_{1,2}^r(x;T_2)}{\partial x}\right|_{x=1} = T_2\left(q_{1|2}\frac{u_1/w_1}{1-T_1+T_1u_1}+(1-q_{1|2})\right)\frac{\langle k_2(k_2-1)(1-T_1+T_1u_1)^{k_1}\rangle}{\langle k_2\rangle}$$
$$= \left(q_{1|2}\frac{u_1/w_1}{1-T_1+T_1u_1}+(1-q_{1|2})\right)R_{0,2}, \quad (30)$$

and the critical transmission rate

$$T_{C,2}^{\text{overlap}} = \frac{1-T_1+T_1u_1}{q_{1|2}u_1/w_1+(1-q_{1|2})(1-T_1+T_1u_1)}\frac{\langle k_2\rangle}{\langle k_2(k_2-1)(1-T_1+T_1u_1)^{k_1}\rangle} = \frac{1-T_1+T_1u_1}{q_{1|2}u_1/w_1+(1-q_{1|2})(1-T_1+T_1u_1)}T_{C,2}. \quad (31)$$

The aforementioned constraints on the degree distribution, given the value of $q_{1|2}$, make it necessary to distinguish the impact of network overlap from correlations in the degree distributions. We can do this by focusing on the multiplicative factors modifying the original values for nonoverlapping networks, $R_{0,2}$ and $T_{C,2}$ in Eqs. (30) and (31), given a joint degree distribution $p(k_1,k_2)$ and overlap $q_{1|2}$. The factor modifying $T_{C,2}$,

$$\frac{1-T_1+T_1u_1}{q_{1|2}u_1/w_1+(1-q_{1|2})(1-T_1+T_1u_1)}, \quad (32)$$

is equal to 1 if $q_{1|2}=0$, and equal to $(1-T_1+T_1u_1)w_1/u_1<1$, if $q_{1|2}=1$ [full overlap, where $u_1\geq w_1$ because $p(k_1,k_2>k_1)=0$]. Between these, it is strictly decreasing with $q_{1|2}$. This means that for a fixed joint degree distribution, invasion for the second process becomes easier if there is more overlap

between the two networks because random nodes outside the GCOC on Γ_1 have a higher tendency to be connected to other nodes outside Γ_1 on Γ_2 if the two networks overlap.

However, it is difficult to separate the impact of overlap from the joint degree distribution, which will usually change, too, with different levels of overlap. This becomes clear, for example, if we consider networks of power-law degree distributions (often called scale-free networks). Such networks, distributed with $p_k \propto k^{-\alpha}$, are often used as proxy for highly heterogeneous networks and their degree distribution has infinite variance for $\alpha<3$, causing the epidemic threshold on single networks to drop to zero because the second moment diverges in Eq. (3). In our case of overlay networks, if the degrees are independently distributed $[p(k_1,k_2) \propto k_1^{-\alpha} k_2^{-\alpha}]$ with $\alpha<3$, and the networks do not overlap ($q_{1|2}=0$), Eq. (31) yields that $T_{C,2}=0$ independent of the value of T_1, so that the first process cannot stop the second one. If, on the other hand, degrees are fully positively correlated $[p(k_1,k_2) \propto \delta_{k_1 k_2} k_1^{-\alpha}]$ with $\alpha<3$ and the networks overlap ($q_{1|2}=1$), we get that $T_{C,2}=1$, so that it becomes impossible for the second process to spread. Just as in Fig. 2, this is because the nodes of high degree are occupied by the first process and prevent the second one from spreading.

C. Partial immunity

Not in all cases that we mentioned in the introduction is the mutual exclusion of the two processes perfect. Sometimes infection with one process gives only partial immunity to the other and a reduced chance of transmission remains. In some cases, the probability of infection is even increased, and one of the two processes makes individuals more susceptible to the other.

We now consider the case where infection with one of the two processes does not make the nodes completely immune to the other, but only modifies their transmission probabilities to σT_2 with $0 \leq \sigma < 1/T_2$. This includes the cases of complete immunity ($\sigma=0$), no interaction ($\sigma=1$), as well as the case when infection with the first process facilitates infection with the second ($\sigma>1$).

In this case, we cannot focus on the residual network anymore, because nodes in the GCOC of Γ_1 can have their edges occupied on Γ_2, too, with modified probability σT_2. To compare this scenario with the ones studied before, we concentrate on clusters on Γ_2 which contain at least one node in the residual network of Γ_1. This is equivalent to saying that our second SIR epidemic starts with an individual unaffected by the first outbreak.

Let us write down $G_{0,2}^{\text{partial}}(x,y)$, the pgf of a random node on the residual network of Γ_1 to be connected to m nodes on Γ_2 which are not in the GCOC of Γ_1 and k_2-m nodes within that GCOC,

$$G_{0,2}^{\text{partial}}(x,y) = \sum_{m,n} \binom{k_2}{m} w_1^m (1-w_1)^{k_2-m} x^m y^{k_2-m} = [y(1-w_1)+xw_1]^{k_2}. \tag{33}$$

Analogously to Eq. (15), we then have

$$G_{0,2}^{\text{partial}}(x,y;T_2) = G_{0,2}^{\text{partial}}(1-T_2+T_2x, 1-\sigma T_2+\sigma T_2 y), \tag{34}$$

as the generating functions for a random node to have s of its m edges on Γ_2 linking to nodes not part of the GCOC in Γ_1 and t of the k_2-m others occupied. The generating function for the total number of edges occupied on Γ_2 starting from a random node on the residual network of Γ_1 on Γ_2 is then

$$G_{0,2}^{\text{partial}}(x;T_2) = G_{0,2}^{\text{partial}}(x,x;T_2). \tag{35}$$

Now we have to distinguish between nodes arrived at following one of the m edges not connected to nodes in the GCOC of Γ_1 and ones arrived at following one of the k_2-m connected to nodes within that GCOC. The probability of a node of degree k_1 on Γ_1 to be part of the GCOC on that network is ϵ_{k_1}, as of Eq. (9). Therefore, the probability of arriving at a node of degree k_2 if it is has been selected following an edge connected to a node in the GCOC of Γ_1 is proportional to $p(k_1,k_2)k_2\epsilon_{k_1}$, and it is proportional to $p(k_1,k_2)k_2(1-\epsilon_{k_1})$ if it has been selected following one of the m others. For the nodes arrived at from a randomly selected node, we then have two different generating functions $G_{1,2}^{+}$ and $G_{1,2}^{-}$ for the number of excess occupied edges, depending on whether that node is in the GCOC on Γ_1 or not, and similarly we can define $H_{1,2}^{+}$ and $H_{1,2}^{-}$ for the generating functions for the size of connected occupied clusters.

The generating function for the occupied cluster size on Γ_2 of a randomly selected node on the residual network of Γ_1 is then

$$H_{0,2}^{\text{partial}}(x;T_2) = G_{0,2}^{\text{partial}}(x;T_2)[H_{1,2}^{-}(x;T_2), H_{1,2}^{+}(x;T_2)]. \tag{36}$$

The generating functions $H_{1,2}^{-}$ and $H_{1,2}^{+}$ fulfill the consistency relations

$$H_{1,2}^{-}(x;T_2) = G_{1,2}^{-}[H_{1,2}^{-}(x;T_2), H_{1,2}^{+}(x;T_2)],$$

$$H_{1,2}^{+}(x;T_2) = G_{1,2}^{+}[H_{1,2}^{-}(x;T_2), H_{1,2}^{+}(x;T_2)]. \tag{37}$$

Using the relations (36) and (37), we can use numerical approximation methods to determine $u_1^{-}=H_{1,2}^{-}(1;T_2)$ and $u_1^{+}=H_{1,2}^{+}(1;T_2)$ and proceed as usually to determine, for example $S_2(T_2)=1-H_0(1;T_2)$ from Eq. (36). While it is possible to derive an analytical expression for $T_{C,2}$ and $R_{0,2}$ in terms of u_1^{-} and u_2^{-} following the same steps as in the cases outlined above, such an endeavor is tedious and provides little opportunity for interpretation, and we prefer here to just present numerical results. Quite intuitively, relaxing the immunity to one process granted by the other one will make it easier for that one to invade the population. This can be seen in Fig. 3, where a second process which would not spread under total exclusion can infect a significant portion of the network if $\sigma>0$.

D. Mutual cross immunity

In the bond percolation model presented here, we can also study a scenario where not only one process is inhibited by not being able to spread to nodes which have any of their

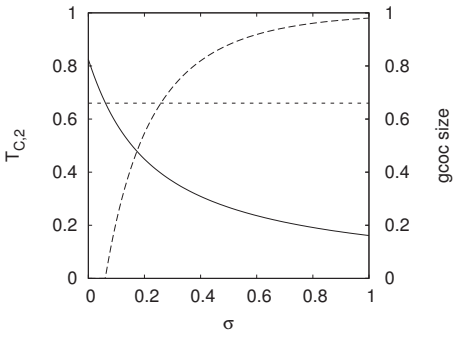

FIG. 3. The critical transmission probability (solid line) and fraction of nodes in the GCOC (long dashes) for a process running on a random network with $T_2=0.66$ after a previously running process with $T_1=0.33$ has granted partial immunity in the sense that nodes in the GCOC of the first process have their edges occupied by the second with probability σT_2.

edges occupied by the other, but also vice versa. That is, we can study situations where both outbreaks have an influence on another. To do so, we search for solutions u_1, u_2 of the system where the GCOC of either network emerges on the residual degree distribution of the other [compare Eq. (14)],

$$G^r_{0,1}(x) = \frac{1}{G_{0,2}(u_2;T_2)} \sum_{k_1}^{\infty} \sum_{k_2}^{\infty} p(k_1,k_2) \times (1 - T_2 + T_2 u_2)^{k_2} (1 - w_2 + w_2 x)^{k_1},$$

$$G^r_{0,2}(x) = \frac{1}{G_{0,1}(u_1;T_1)} \sum_{k_1}^{\infty} \sum_{k_2}^{\infty} p(k_1,k_2) \times (1 - T_1 + T_1 u_1)^{k_1} (1 - w_1 + w_1 x)^{k_2}, \quad (38)$$

with w_2 the probability of finding a node in the GCOC of Γ_2 by following an edge on Γ_1, analogously to w_1 in Eq. (11). We can solve system (38) numerically under various scenarios of joint degree distribution, partial immunity, and network overlap. Remembering that we assumed large outbreaks always to happen if they are possible, the solutions u_1, u_2 then correspond to the probabilities for nodes not to be part of outbreaks of SIR epidemics on Γ_1 and Γ_2, respectively, where both these epidemics interact with each other, and solutions where both $u_1 < 1$ and $u_2 < 1$ indicate situations where the two processes can coexist.

Depending on the networks and the values of the transmission probabilities T_1 and T_2 with regard to their critical probabilities $T_{1,C}$ and $T_{2,C}$, we find qualitatively different types of possible solutions to Eq. (38). Generally, either of the two processes can always form a GCOC if its transmission probability is greater than its critical transmission probability $T_{1/2,C}$ as determined from the marginal degree distribution using Eq. (3), i.e., ignoring the presence of the other network. Only if both T_1 and T_2 are below their critical values, there can be completely nonepidemic solutions, i.e., solutions where neither of the two processes form a GCOC. If, on the other hand, one of them is above the critical value whereas the other is not, only the one with its transmission probability $T_{1/2}$ greater than its critical transmission probability $T_{1/2,C}$ will form a GCOC.

The most interesting situation arises if both T_1 and T_2 are greater than their respective critical values. In that case, we observe three different cases, depending on the transmission probabilities and network properties. In one case there is only one solution to system (38) with one of the two processes forming a GCOC whereas the other one remains confined to small clusters. A second case again yields only one solution, but this time the two processes can coexist and both form a GCOC. This corresponds to a situation of pure coexistence, where, even if one of the two processes forms a GCOC, there is always still enough of the network left in the residual network for the other one to form a GCOC, too. Lastly, there can be situations where there are three solutions to Eq. (38), one of coexistence and two more with either of the two processes forming a GCOC exclusively, the other one being confined to small clusters. In that case, if one process forms a GCOC without interference from the other one, there will not be enough of the network left in the residual network for the other one to form a GCOC, too. However, there is also the possibility that both processes limit each other to leave enough of the network for both to form a GCOC.

How does the analogy with the SIR model work in this case? If both processes do not form a GCOC, no large outbreak happens on either SIR process, whereas if only one can form a GCOC, the corresponding SIR process forms a large outbreak. Where there is exactly one solution to Eq. (38) with a GCOC both on Γ_1 and Γ_2, both SIR epidemics cause a large outbreak, and they coexist on the network. Lastly, where there are three solutions to Eq. (38), three different scenarios can occur. Either one of the two processes causes a large outbreak and occupies a sufficient part of the network to prevent the other one from spreading, or both cause a large but smaller outbreak, leaving a large enough part of the network for the other process to spread, so that the two coexist. The percolation model does not tell us anything about which of these situations will actually occur in a single run of a stochastic SIR model and this will partly depend on the relative velocities of the two epidemics, more specifically on whether one of them spreads fast enough with respect to the other to occupy a sufficient fraction of the network to suppress it.

Figures 4 and 5 show regions of possible coexistence for different joint degree distributions, where both networks have the same marginal degree distribution. In that case, we see that coexistence is always possible if $T_1 \approx T_2$, and they are both greater than their critical transmission probability T_C. For independently distributed degrees and complete mutual immunity, with increasing heterogeneity there is a growing region around $T_{1/2} \gtrsim T_C$ in which there is always coexistence, i.e., both processes always form a GCOC, although this is not the case if the degrees of individual nodes on the network are positively correlated.

When, on the other hand, T_1 or T_2, approach one, and both processes grant complete immunity to each other, there is always a solution with either of the two processes occupying such a large part of the network that the other cannot spread,

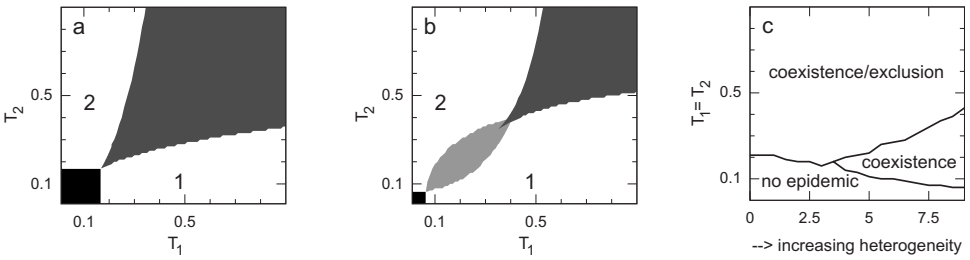

FIG. 4. (a) and (b) Regions of possible coexistence for two epidemics spreading with transmission probabilities T_1 and T_2 and complete mutual immunity on nonoverlapping overlay networks with independently distributed degrees of average degree $\langle k_1\rangle=\langle k_2\rangle=6$ with standard deviation (a) $s=2.5$ and $s=8$ (b). Regions where none of the two processes can form a GCOC are black and ones where only one can form a GCOC are white, with the number inside the area indicating which of the two processes is able to do so. In regions of dark gray, there can be coexistence, but also exclusion of either process by the other, whereas in areas or light gray, there is always coexistence. (c) Regions of possible coexistence for $T_1=T_2$ as a function of standard deviation in node degree for complete mutual immunity on nonoverlapping overlay networks with independently distributed degrees of average degree $\langle k_1\rangle=\langle k_2\rangle=6$.

in addition to a potentially existing solution of coexistence. This changes if the immunity granted mutually is only partial, where we see that there is always coexistence for T_1 and T_2 sufficiently large. Only if both networks have little heterogeneity (or if they are regular), and T_1 and T_2 are close to T_C there can be solutions where either of the two processes can form a GCOC large enough to prevent the other one from spreading, in addition to the possibility of coexistence.

IV. CONCLUSIONS

In conclusion, we have developed a framework for the study of interacting epidemics on overlay networks of arbitrary joint degree distribution, amount of overlap and strength of immunity. If one process spreads to grant protection to another one spreading subsequently, such as awareness of a disease might do to a disease, or a pathogen to a similar one (see the contemporary work of [22]), the protection is more effective if the degrees of nodes on the two networks are positively correlated. In that case, the protective effect is stronger if there is greater variance in the degree distribution. If, on the other hand the degrees are uncorrelated or even negatively correlated, increasing heterogeneity makes it more and more difficult to contain the second epidemic.

This generalizes the notion of a coexistence threshold introduced in [15] to situations where the two pathogens do not spread on the same network or where the network changes between subsequent outbreaks, and where they do not have the same transmission probability. If, on the other hand, we do apply these conditions to our model, the coexistence threshold emerges as a special case.

If both processes mutually affect each other, the parameter regions of possible coexistence can be quite small, especially if the transmission probabilities are close to the respective thresholds. If both grant only partial immunity to each other, this effect becomes even stronger. If the transmission probabilities are relatively high, on the other hand, partial immunity excludes dominance of one process because either one can always establish itself even with the other one already occupying parts of the system.

The method applied here is static in nature and fails to predict the dynamical interaction of two processes in competition. This is done more naturally using models based on differential equations, as done, for example, in [7]. These models, however, often fail to account for the influence of population structure. To capture dynamical features of the interaction of two processes competing over some part of a network, other approaches such as the generation-based method recently introduced in [23] might hold more promise. Still, the results presented here show the wide range of possible impacts of large-scale network structure on the interaction between multiple spreading agents, highlighting once more that network structure and interplay can affect the patterns of concurrent spread of two processes, whether it be rumors, pathogens or others.

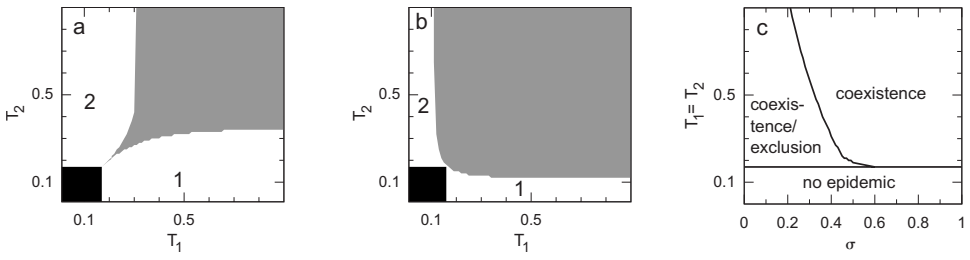

FIG. 5. (a) and (b) The effect of partial immunity $\sigma=0.5$ (a) and mutual facilitation $\sigma=1.5$ (b) on the regions of possible coexistence for the network used in Fig. 4(a) (average degree $\langle k_1\rangle=\langle k_2\rangle=6$ with standard deviation $s=2.5$). (c) Regions of possible coexistence for $T_1=T_2$ as a function of partial immunity σ in the same network.

ACKNOWLEDGMENTS

This research was funded by the UK Engineering and Physical Sciences Research Council through standard research grant number EP/D002249/1.

APPENDIX

The overlay networks used to create the figures were created in the following way: single networks of $N=10\,000$ nodes and average degree $\langle k\rangle=6$ and standard deviation in the degree $s=0$, $s=2.5$, and $s=9$ were created using the algorithms of [24–26], respectively. Single networks with intermediate heterogeneity were generated by randomizing a fraction of the edges in these networks.

Joint degree distributions for overlay networks with independently distributed degrees, if both networks are to have the same marginal degree distribution, can then be generated by setting $p(k_1,k_2)=p(k_1)p(k_2)$, where $p(k)$ is the degree distribution of the single network. The corresponding overlay networks for simulations can be generated by starting from the single network to represent Γ_1 and randomly reshuffling the vertices of a second single network representing Γ_2 before adding its edges.

To generate overlay networks with positive or negative degree correlations, we used a rewiring method based on the Metropolis algorithm [see, e.g., [27,28], for similar approaches] with a Hamiltonian defined by

$$\mathcal{H}(\Gamma)=-J\sum_{v=1}^{N}k_{1,v}k_{2,v}, \tag{A1}$$

where Γ denotes the overlay network, and the sum is over all vertices and $k_{1,v}$ and $k_{2,v}$ are the degrees of vertex v on the network Γ_1 and Γ_2, respectively. A positive (negative) value of the constant J will favor positive (negative) degree correlations in the edge swap algorithm described in the following.

Starting from the overlay network with independently distributed degrees, we select two distinct random vertices v_1 and v_2, and propose an edge swap between the two, in the sense that v_1 receives all links v_2 has on Γ_2 and vice versa, subject to not creating overlapping edges, but they both keep their links on Γ_1. Denoting the original overlay network with Γ and the network which would result from the edge swap with Γ', the swap is then accepted with probability

$$\min(1,\exp\{-[\mathcal{H}(\Gamma')-\mathcal{H}(\Gamma)]\}). \tag{A2}$$

If we measure the degree correlation using the Pearson product-moment correlation coefficient [29]

$$r=\frac{\sum_{v}(k_{1,v}-\langle k_1\rangle)(k_{2,v}\langle k_2\rangle)}{\sqrt{\left[\sum_{v}(k_{1,v}-\langle k_1\rangle)^2\sum_{v}(k_{2,v}-\langle k_2\rangle)^2\right]}}, \tag{A3}$$

we can stop rewiring once we have obtained the desired correlation r. In the text, positive correlation indicates $r=0.7$ and negative correlation $r=-0.7$.

To approximate T_C from the dynamical SIR model on random networks of size $N=10\,000$ in Fig. 2, we ran individual-based simulations on the networks created using the method outline above. In these simulations we first ran an outbreak of the first epidemic with transmission probability T_1, and removed all nodes infected in the outbreak from the population. We then ran the second epidemic on the remaining network and tracked the average number of individuals infected by infected individuals of the second generation, i.e., those who were infected by the very first infected case of the second epidemic. This average number is our $R_{0,2}$. By calculating this for different values of T_2 we got a relation between T_2 and $R_{0,2}$ which should be linear. We then fitted a straight line to this relation and took the value of T_2 which yielded $R_{0,2}=1$ to be the critical transmission probability $T_{C,2}$.

[1] M. E. J. Newman, Phys. Rev. E **66**, 016128 (2002).
[2] M. J. Keeling and K. T. D. Eames, J. R. Soc., Interface **2**, 295 (2005).
[3] R. M. May, Trends Ecol. Evol. **21**, 394 (2006).
[4] L. Elveback, J. P. Fox, and A. Varma, Am. J. Trop. Med. Hyg. **80**, 356 (1964).
[5] I. W. Saunders, J. Math. Biol. **11**, 311 (1981).
[6] L. J. Abu-Raddad, P. Patnaik, and J. G. Kublin, Science **314**, 1603 (2006).
[7] D. A. Vasco, H. J. Wearing, and P. Rohani, J. Theor. Biol. **245**, 9 (2007).
[8] J. R. Gog and B. T. Grenfell, Proc. Natl. Acad. Sci. U.S.A. **99**, 17209 (2002).
[9] L. J. Abu-Raddad, B. I. S. van der Ventel, and N. M. Ferguson, Phys. Rev. Lett. **100**, 168102 (2008).
[10] S. Funk, E. Gilad, C. Watkins, and V. A. A. Jansen, Proc. Natl. Acad. Sci. U.S.A. **106**, 6872 (2009).
[11] S. Funk, E. Gilad, and V. A. A. Jansen, J. Theor. Biol. **264**, 501 (2010).
[12] K. T. D. Eames, J. R. Soc., Interface **6**, 811 (2009).
[13] P. Wang, M. C. González, C. A. Hidalgo, and A.-L. Barabási, Science **324**, 1071 (2009).
[14] C. Peikari, " Fighting fire with fire: designing a "good" computer virus," (2004), *InformIT*, URL: http://www.informit.com/articles/article.aspx?p=337309, archived at http://www.webcitation.org/5kXXvHXfI on Oct 15, 2009.
[15] M. E. J. Newman, Phys. Rev. Lett. **95**, 108701 (2005).
[16] R. M. Anderson, G. F. Medley, R. M. May, and A. M. Johnson, IMA J. Math. Appl. Med. Biol. **3**, 229 (1986).
[17] H. Andersson, Math. Biosci. **140**, 79 (1997).
[18] O. Diekmann and J. A. P. Heesterbeek, *Mathematical Epidemiology of Infectious Diseases* (Wiley, Chichester, 2000).
[19] L. A. Meyers, Bull. Am. Math. Soc. **44** (1), 63 (2007).
[20] R. Bartoszyński, in *Proceedings of the Fifth Berkeley Sympo-*

sium on Mathematical Statistics and Probability Vol. 4 (University of California Press, Berkeley, 1967) pp. 259–269.
[21] H. S. Wilf, *Generatingfunctionology* (Academic, London, 1990).
[22] S. Bansal, B. Pourbohloul, N. Hupert, B. Grenfell, L. A. Meyers, PLos One **5** (2), e9360 (2010); S. Bansal and L. A. Meyers, e-print arXiv:0910.2008 (2009).
[23] P.-A. Noël, B. Davoudi, R. C. Brunham, L. J. Dubé, and B. Pourbohloul, Phys. Rev. E **79**, 026101 (2009).
[24] J. Kim and V. Vu, Combinatorica **26**, 683 (2006).
[25] P. Erdős and A. Rényi, Publ. Math. (Debrecen) **6**, 290 (1959).
[26] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[27] C. Biely and S. Thurner, Phys. Rev. E **74**, 066116 (2006).
[28] J. D. Noh, Phys. Rev. E **76**, 026116 (2007).
[29] J. Rodgers and W. Nicewander, Am. Stat. **42**, 59 (1988).