Cocirculation of infectious diseases on networks

Joel C. Miller

Department of Mathematics and Department of Biology, Penn State University, University Park, Pennsylvania 16802, USA

We consider multiple diseases spreading in a static configuration model network. We make standard assumptions that infection transmits from neighbor to neighbor at a disease-specific rate and infected individuals recover at a disease-specific rate. Infection by one disease confers immediate and permanent immunity to infection by any disease. Under these assumptions, we find a simple, low-dimensional ordinary differential equations model which captures the global dynamics of the infection. The dynamics depend strongly on initial conditions. Although we motivate this Rapid Communication with infectious disease, the model may be adapted to the spread of other infectious agents such as competing political beliefs, or adoption of new technologies if these are influenced by contacts. As an example, we demonstrate how to model an infectious disease which can be prevented by a behavior change.

Introduction. Many important dynamical systems involve the spread of some process through a network [1–3]. Our ability to understand the systems is complicated by a lack of analytic models, limiting our understanding of the dynamics. Particularly when competing processes are occurring in the network, the final outcome may be highly dependent on the intermediate dynamics, so the lack of an analytic model prevents us from being able to predict even the long-term equilibrium state of a system. In this Rapid Communication we study the simultaneous spread of two competing diseases in a configuration model network. Although we focus on disease, other competing "infectious" processes such as a change in behavior in response to a disease [4], spread of beliefs in a voter model [2], and "viral marketing" of competing technologies [5] have been studied, and the approach introduced here can be adapted to these applications.

In this Rapid Communication we derive a low-dimensional system of equations capturing the dynamics of competing diseases spreading simultaneously in a configuration model network. We apply the model in investigating possible outcomes of cocirculating diseases. Prior studies have extensively analyzed the effect of network structure such as degree distribution [6–9] and heterogeneities in infectiousness and/or susceptibility [10–13] on disease spread. Recent work gives insight into the role of partnership duration [14–17]. Other investigations focus on the role of clustering [13,18–25], with limited predictive success.

Models of interacting diseases typically neglect network structure (e.g., [26,27], and many others). Until recently, models of a single disease spreading through a network have relied on systems of many equations [28] or been restricted to final size calculations under the assumption of an asymptotically small initial fraction infected [6]. Extending these approaches to competing diseases [29–31] does not allow us to measure the effect of dynamic interactions, and so results are limited to special cases in which these interactions are not important, such as when one disease spreads before the other. Some recent work allows us to investigate simultaneous disease spread in "overlay networks", but with a potentially unbounded number of equations [32].

The method we introduce allows us to capture the dynamic interactions of two diseases spreading in a configuration model network and allows for arbitrarily large initial conditions. It can be easily adapted to more than two diseases. We validate the system by comparison with simulation. Using our equations, we are able to identify the scalings which separate different regimes. We discuss these regimes and introduce possible generalizations [33].

The basic model. We assume that two diseases spread in a configuration model network [34] (also called a Molloy-Reed network [35]) with the degree distribution given by $P(k)$. For disease 1 transmission along an edge has rate β_1 and recovery of infected individuals has rate γ_1. For disease 2 the rates are β_2 and γ_2. A node infected by either disease gains immunity to any further infection.

Our approach is similar to that of [36] and is based on [15]. We will focus our attention on a *test individual u*, a randomly chosen individual in the population [37]. We assume that the aggregate population-scale spread of the diseases is deterministic. Under these assumptions, the probability the test individual has a given infection status equals the proportion of the population with that status. Thus by calculating the probability a test individual has a given status, we immediately know the proportion of the population with that status.

We make one change to the test individual u: we prevent it from causing infections. This keeps the status of its partners independent of one another without affecting its own status, and so it has no effect on our calculations of the proportion of the population in each state. An alternate argument for why this change has no impact is that we have assumed the dynamics are deterministic, while the timing of when (or even if) u is infected is a random variable. Thus the infections u would cause cannot have any macroscopic impact on the disease dynamics, and this modification of u has no effect at the population scale. This is similar to the concept of a thermal bath because the population-scale dynamics are independent of what u does. It is directly analogous to the "price taker" assumption of economics for which a firm does not produce enough of a commodity to alter the market price, and so we can ignore its impact on the market when we calculate the market's impact on it. In our case, because we know the test individual's local effect is unimportant at the population scale, we are able to remove its local effect without altering the population-scale dynamics. Thus we can ignore the impact

of u on its environment while calculating the impact of the environment on u.

We take $t = t_0$ to be our "initial time.' In practice this may correspond to the time of introduction of a disease if enough individuals are initially infected, or it corresponds to a later time at which enough infection is present that the spread is deterministic. We choose a test individual u randomly from the population (it may have any status). We let v be a random neighbor of u which had not transmitted to u by t_0. We assume that the status of v is independent of the degree of u. We define $\theta(t)$ to be the probability that at time t, v has not transmitted to u. The probability u is susceptible at time t is

$$S(t) = \psi(\theta(t)) = \sum_k P(k)S(k,t_0)\theta(t)^k,$$

where $S(k,t_0)$ is the probability an individual of degree k is susceptible at $t = t_0$. We take I_1 and I_2 (respectively, R_1 and R_2) to be the probabilities that u is infected with (respectively, has recovered from) the corresponding disease.

To calculate the change of θ, we must know more about the probability v is in any given state. We define $\phi_S(t)$ to be the probability v is susceptible; $\phi_{I,1}(t)$ and $\phi_{I,2}(t)$ to be the probabilities that v is infected, has not transmitted to u and is infected by disease 1 or 2 respectively; and ϕ_R to be the probability v is recovered but did not transmit (we do not need to distinguish which disease infected v). Then $\theta = \phi_S + \phi_{I,1} + \phi_{I,2} + \phi_R$.

We calculate ϕ_S similarly to S. If v is initially susceptible we find the probability it has degree k by counting all edges of initially susceptible individuals of degree k: $NkP(k)S(k,t_0)$ and dividing by the number of all edges of initially susceptible individuals $\sum_k NkP(k)S(k,t_0)$ (N is population size). If v has degree k and was initially susceptible, the probability v is still susceptible is θ^{k-1} (because u is prevented from transmitting to v). This leads to the conclusion that v is susceptible with probability $\phi_S(t) = \phi_S(t_0)\frac{\sum_k kP(k)S(k,t_0)\theta^{k-1}}{\sum_k kP(k)S(k,t_0)} = \phi_S(t_0)\frac{\psi'(\theta)}{\psi'(1)}$.

Figure 1 gives flow diagrams which yield our equations. Each box represents a compartment, and arrow labels represent probability flux from one compartment to another. The fluxes from the I compartments to the R compartments represent recovery of u. The fluxes from the ϕ_I compartments to the ϕ_R compartments (respectively, $1 - \theta$) represent flux due to recovery of v prior to transmitting to u (respectively, transmission to u prior to recovery of v). The fluxes from S and ϕ_S are found by differentiation of S and ϕ_S in time,

using $\dot{\theta} = -\beta_1\phi_{I,1} - \beta_2\phi_{I,2}$, and assigning the appropriate proportion to the appropriate compartment.

From the diagram, we find

$$\dot{\theta} = -\beta_1\phi_{I,1} - \beta_2\phi_{I,2}, \tag{1}$$

$$\dot{\phi}_{I,m} = -(\beta_m + \gamma_m)\phi_{I,m} + \beta_m\phi_{I,m}\phi_S(t_0)\frac{\psi''(\theta)}{\psi'(1)}, \tag{2}$$

$$S = \psi(\theta), \tag{3}$$

$$\dot{I}_m = \beta_m\phi_{I,m}\psi'(\theta) - \gamma_m I_m, \tag{4}$$

$$\dot{R}_m = \gamma_m I_m. \tag{5}$$

The subscript m takes the values 1 and 2. The single-disease, small initial condition limit of these equations has been proved exact [38]. These equations capture the fact that disconnected components are safe from outside introduction.

Sequential introduction. As an example, we consider two diseases spreading in a network of 10^6 individuals with Poisson degree distribution of mean 5. For both, $\gamma = 1$ but $\beta_1 = 0.5$ and $\beta_2 = 1$. In simulations, we introduce disease 1 into 30 random individuals at $t = 0$, and disease 2 into 30 random individuals at $t = 5.25$.

Our deterministic equations do not apply while either disease has a small number of infections. We take observations at $t_0 = 4$ (after the first, but before the second disease) and $t_0 = 7$ (after both are established) to initialize our equations. Figure 2 compares calculation with simulation. Comparing the $t_0 = 4$ calculation with the $t_0 = 7$ calculation shows the effect of the second disease.

Simultaneous introduction. We now consider the simultaneous introduction of two diseases and assume that the initial numbers infected are large enough that the dynamics are deterministic. In our example, we take $\beta_1 = 1.2$, $\gamma_1 = 4$, $\beta_2 = 0.2$, and $\gamma_2 = 0.25$. Disease 1 tends to spread more quickly, but disease 2 has a higher probability of transmission prior to recovery. At $t = 0$, we randomly selected a proportion ρ_1 to be infected with disease 1, and a proportion ρ_2 to be infected with disease 2. This gives $S(k,0) = 1 - \rho_1 - \rho_2$ for all k, $I_1 = \rho_1$, $I_2 = \rho_2$, $\phi_{I,1} = \rho_1$, and $\phi_{I,2} = \rho_2$, with no recovered individuals. In our population, the degree of each node is assigned uniformly from the integers 1–9. We use our equations to calculate the final proportion infected by each disease, shown in Fig. 3.

There are several distinct regimes we can identify in Fig. 3. If ρ_2 is $O(1)$ and ρ_1 small, or if ρ_1 is $O(1)$ and ρ_2 small, the disease with the large initial condition spreads and effectively infects everyone simply because a large fraction is initially infected. The other disease cannot spread in the "residual network" left behind. If neither ρ_1 or ρ_2 is initially large, other regimes are seen. To analyze them, we first note that when both diseases are small, they grow at exponential rates $r_1 = -(\beta_1 + \gamma_1) + \beta_1\psi''(1)/\psi'(1) = 2$ and $r_2 = -(\beta_2 + \gamma_2) + \beta_2\psi''(1)/\psi'(1) = 0.75$.

In the "overlapping epidemic" regime, the slower growing disease 2 begins with a head start. The size of the head start scales so that the two epidemics become large at the same time. The value of $\ln I_2 - (r_2/r_1)\ln I_1$ is constant during the exponential growth phase. For given $C = \ln\rho_2 - (r_2/r_1)\ln\rho_1$,
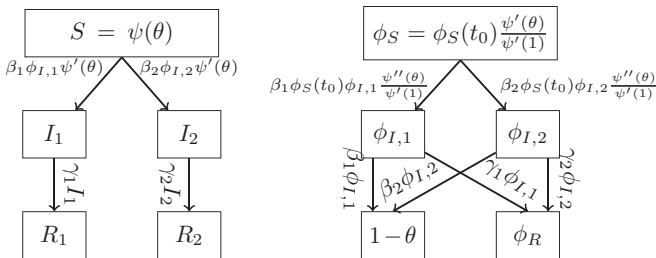

FIG. 1. (Left) Flow diagram for the probabilities the test individual has each status. (Right) Flow diagram for the probabilities a partner of the test individual has each status.
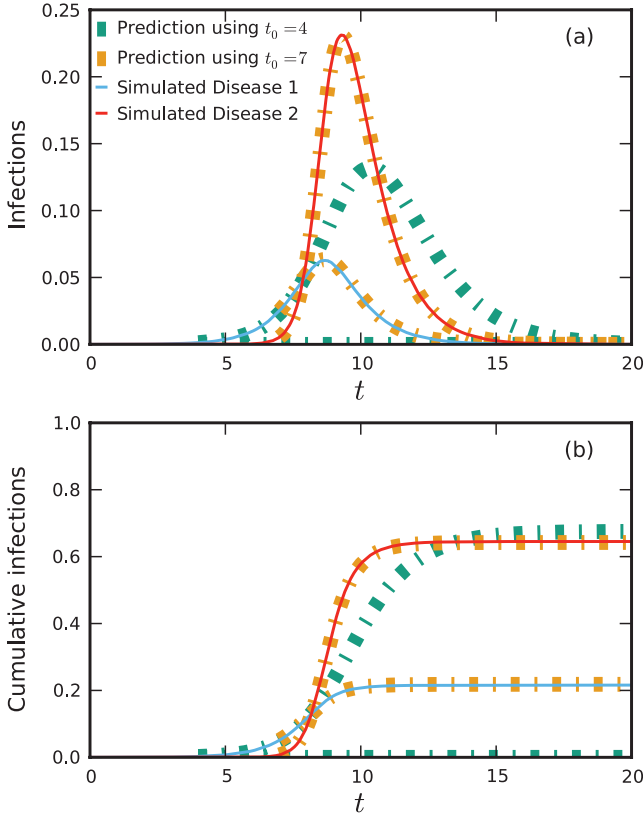
FIG. 2. (Color online) The spread of two diseases in a population of size 10^6 with a Poisson degree distribution of mean 5. The first disease is introduced with 30 cases at $t = 0$ and the second with 30 cases at $t = 5.25$. The second strain is more infectious. Predictions (dashed) are calculated from observations at $t = 4$ (before the second disease's introduction) and $t = 7$ (shortly after). Model and simulation agree well.

FIG. 3. (Color online) We take a network with the degree of each node chosen uniformly from 1 up to 9. A proportion ρ_1 of the population begins infected with disease 1 and a proportion ρ_2 begins infected with disease 2 with $\beta_1 = 1.2$, $\gamma_1 = 4$, $\beta_2 = 0.2$, and $\gamma_2 = 0.25$. Disease 1 spreads more quickly, but has a smaller per-edge transmission probability. (a) Final proportion infected with disease 1. (b) Final proportion infected with disease 2. The solid lines denote the estimated upper and lower bounds of the overlapping epidemic regime.

the behavior is universal. The diseases grow independently until the exponential growth phase ends. The amount of each disease at this time is determined by C. We can estimate bounds on the regime by crudely assuming exponential growth continues forever. There is some value of $C_{\min} = \ln 0.0025 - (r_2/r_1) \ln 1 = \ln 0.0025 \approx -6$ that corresponds to $\rho_1 = 1$ and $\rho_2 = 0.0025$, which means that the slower growing disease would affect less than 1% of the population by the time the faster growing disease has fully established itself in this approximation. Similarly we take some $C_{\max} = \ln 1 - (r_2/r_1) \ln 0.05 = -(r_2/r_1) \ln 0.05 \approx 3r_2/r_1$ corresponding to $\rho_1 = 0.05$ and $\rho_2 = 1$, which means that the slower growing disease will have fully burned through the population when the faster growing disease is only affecting 5% of the population in this approximation. For $C < C_{\min}$, I_1 becomes large well before I_2. For $C > C_{\max}$, I_2 becomes large well before I_1. Between these values, the epidemics become large at similar times and interact dynamically.[1]

There are two "nonoverlapping epidemic" regimes. If $C < C_{\min}$, disease 1 becomes large and has an epidemic
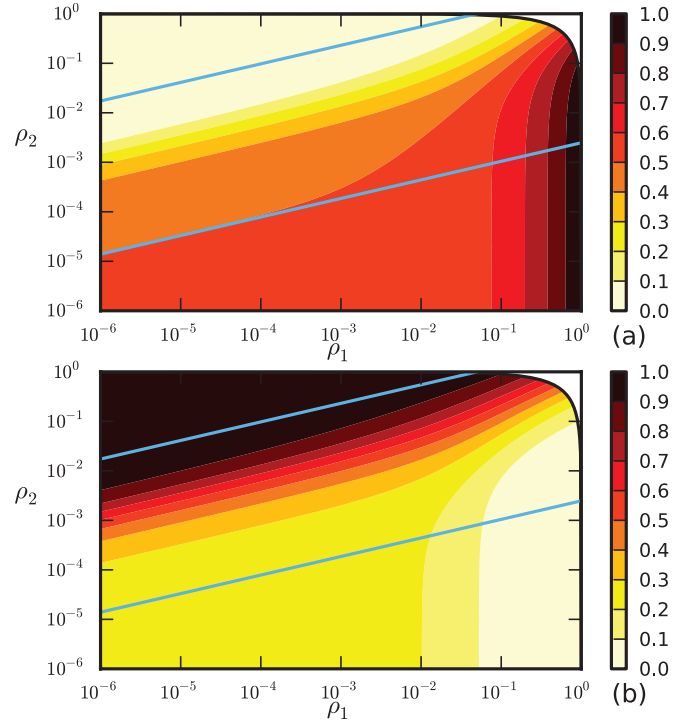
while disease 2 is still exponentially small. If $C > C_{\max}$, disease 2 has an epidemic while disease 1 is still exponentially small. Once one epidemic has finished, it may be possible for the remaining disease to spread in the residual network. We can derive the threshold condition: For simplicity we assume disease 1 spreads first. While it spreads, $\phi_{I,2}$ remains negligible. By looking at the relative fluxes out of $\phi_{I,1}$ we conclude that after disease 1 has completed its spread but before disease 2 is significant, $\phi_R = \gamma_1(1 - \theta)/\beta_1$. We can assume $\phi_S(0) = 1$ and $\phi_S(t) = \psi'(\theta)/\psi'(1)$. Since $\phi_{I,1}$ and $\phi_{I,2}$ are effectively zero, we conclude $\theta = \phi_S + \phi_R = \psi'(\theta)/\psi'(1) + \gamma_1(1 - \theta)/\beta_1$, yielding

$$\theta = \frac{\beta_1}{\beta_1 + \gamma_1}\left(\gamma_1 + \frac{\psi'(\theta)}{\psi'(1)}\right).$$

Then $R_1 = 1 - \psi(\theta)$. For disease 2 to spread, we require $\dot{\phi}_{I,2} > 0$, which implies $\beta_2/(\beta_2 + \gamma_2) > \psi'(1)/\psi''(\theta)$, where θ comes from the above equation. A similar calculation would lead to a final size for R_2, so in this case we can calculate the final outcomes of the epidemics without calculating the dynamics. We note that the threshold we have found requires that $\beta_2/(\beta_2 + \gamma_2)$ be greater than $\beta_1/(\beta_1 + \gamma_1)$ by a nonzero amount. It is not enough for the second disease's transmission probability to simply be larger than the first, it must be well above that of the first. This is because even once disease 1 peaks and begins to decrease, θ will continue to decrease further,

---

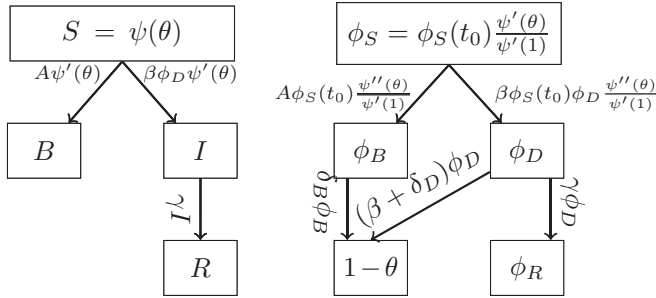[1]A more detailed derivation of these boundaries is in the Supplemental Material [33].

FIG. 4. The flow diagram for disease spread with behavior change. Here $A = (\delta_B \phi_B + \delta_D \phi_D)$.
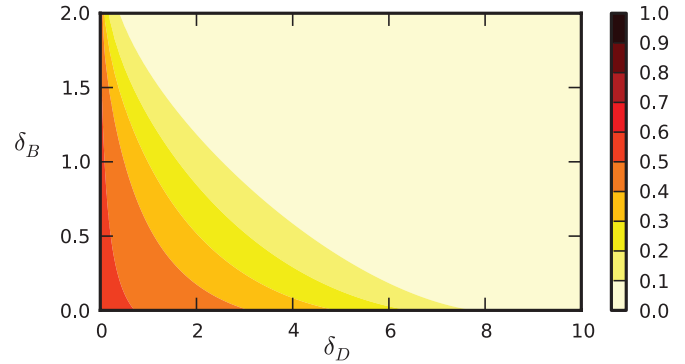

FIG. 5. (Color online) Computed final proportion infected for epidemics with behavior change. We use networks of the same structure as in Fig. 3. The disease spreads with $\beta = 2$, $\gamma = 1$. The values of δ_B and δ_D are varied. For the initial condition, no individuals have changed behavior and a proportion 10^{-6} is initially infected randomly.

so disease 2 encounters a population that is well below the threshold for disease 1. The threshold condition for disease 2 to invade has been identified previously: [29] derived it under the assumption of a second disease introduced after the first disease had spread, and [30] derived it in the special case that the system was in the nonoverlapping epidemic regime for which the faster growing disease would "win."

Thus our analysis shows that if two diseases are introduced at small levels into the network, then the possible regimes can be understood by looking at the exponential growth rates r_1 and r_2 of the diseases. Without loss of generality, we can assume $r_1 \geqslant r_2$: The first disease grows faster. If disease 2 has a sufficiently large head start, there will be a nonoverlapping epidemic regime: The population will experience an epidemic of disease 2 unaffected by disease 1. If the infectiousness of disease 1 is large enough it will cause its own epidemic after disease 2 has finished. We can calculate the final size of each epidemic without requiring the full dynamic calculation. If we take a smaller head start for disease 2, there is an overlapping epidemic regime in which the two diseases produce interacting epidemics. To calculate the dynamics of these epidemics, we require the dynamic equations (1)–(5). The possible final sizes depend on the details of the interactions, and there appears to be no simple expression for the final size. If the head start for disease 2 shrinks further, we enter another nonoverlapping epidemic regime in which disease 1 has the first epidemic. Again, it is possible for disease 2 to later have an epidemic if its transmission probability is sufficiently larger than that of disease 1. At most one of the nonoverlapping epidemic regimes can have epidemics for both diseases.

Generalizations. This model may be adapted to other infectious agents. As an example of its flexibility we consider a disease which can be avoided through behavior modification. We assume contact with an infected individual transmits infection at rate β. However, we allow that if u is in contact with an infected individual, then at rate δ_D u changes its behavior. If u is in contact with an individual who has changed behavior, then u changes behavior at rate δ_B. This leads to the flow diagram in Fig. 4.

In Fig. 5, we show how behavior change modifies epidemic outcomes, using B to denote individuals whose behavior has changed. If the disease is introduced in a small number of individuals and behavior change spreads sufficiently faster than the disease, then the behavior change prevents the disease from having a large-scale epidemic.

Summary. We have introduced an analytic model that calculates the simultaneous spread of two infectious diseases in a configuration model network. Our model is low dimensional regardless of the degree distribution. Using this model, we are able to calculate the effect of interactions between diseases in regimes that are inaccessible to analytic theories that do not include dynamics. This model can easily be generalized to a range of other infectious processes. We have shown its application to behavior changes in response to a disease.

[1] D. B. Larremore, W. L. Shew, and J. G. Restrepo, Phys. Rev. Lett. **106**, 058101 (2011).

[2] R. Durrett, J. P. Gleeson, A. L. Lloyd, P. J. Mucha, F. Shi, D. Sivakoff, J. E. S. Socolar, and C. Varghese, Proc. Natl. Acad. Sci. USA **109**, 3682 (2012).

[3] R. Pastor-Satorras and A. Vespignani, Phys. Rev. E **63**, 066117 (2001).

[4] S. Funk, E. Gilad, C. Watkins, and V. A. A. Jansen, Proc. Natl. Acad. Sci. USA **106**, 6872 (2009).

[5] S. Bharathi, D. Kempe, and M. Salek, in *Internet and Network Economics*, Lecture Notes in Computer Science Vol. 4858, edited by X. Deng and F. Graham (Springer, Berlin, Heidelberg, 2007), pp. 306–311.

[6] M. E. J. Newman, Phys. Rev. E **66**, 016128 (2002).

[7] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).

[8] M. Boguñá, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. Lett. **90**, 028701 (2003).

[9] R. M. May and A. L. Lloyd, Phys. Rev. E **64**, 066112 (2001).

[10] J. C. Miller, Phys. Rev. E **76**, 010101(R) (2007).

[11] J. C. Miller, J. Appl. Probab. **45**, 498 (2008).

[12] E. Kenah and J. M. Robins, Phys. Rev. E **76**, 036113 (2007).

[13] P. Trapman, Theor. Popul. Biol. **71**, 160 (2007).

[14] E. M. Volz and L. A. Meyers, Proc. R. Soc. London, Ser. B **274**, 2925 (2007).

[15] J. C. Miller, A. C. Slim, and E. M. Volz, J. R. Soc., Interface **9**, 890 (2012).

[16] J. C. Miller and E. M. Volz, arXiv:1106.6344.

[17] J. C. Miller and E. M. Volz, J. Math. Biol. (2012), doi:10.1007/s00285-012-0572-3.

[18] J. C. Miller, J. R. Soc., Interface **6**, 1121 (2009).

[19] J. C. Miller, Phys. Rev. E **80**, 020901(R) (2009).

[20] M. E. J. Newman, Phys. Rev. Lett. **103**, 058701 (2009).

[21] J. P. Gleeson, S. Melnik, and A. Hackett, Phys. Rev. E **81**, 066114 (2010).

[22] S. Melnik, A. Hackett, M. A. Porter, P. J. Mucha, and J. P. Gleeson, Phys. Rev. E **83**, 036112 (2011).

[23] E. M. Volz, J. C. Miller, A. Galvani, and L. A. Meyers, PLoS Comput. Biol. **7**, e1002042 (2011).

[24] T. House and M. Keeling, J. R. Soc., Interface **8**, 67 (2010).

[25] K. T. D. Eames, Theor. Popul. Biol. **73**, 104 (2008).

[26] V. Andreasen, J. Lin, and S. A. Levin, J. Math. Biol. **35**, 825 (1997).

[27] S. Cobey and M. Pascual, J. Theor. Biol. **270**, 80 (2011).

[28] K. D. Eames and M. Keeling, Proc. Natl. Acad. Sci. USA **99**, 13330 (2002).

[29] M. E. J. Newman, Phys. Rev. Lett. **95**, 108701 (2005).

[30] B. Karrer and M. E. J. Newman, Phys. Rev. E **84**, 036106 (2011).

[31] S. Funk and V. A. A. Jansen, Phys. Rev. E **81**, 036118 (2010).

[32] V. Marceau, P. A. Noël, L. Hébert-Dufresne, A. Allard, and L. J. Dubé, Phys. Rev. E **84**, 026105 (2011).

[33] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.87.060801 for more mathematical details of the derivations, regime analysis, and investigation of behavior change.

[34] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).

[35] M. Molloy and B. Reed, Random Struct. Algorithms **6**, 161 (1995).

[36] B. Karrer and M. E. J. Newman, Phys. Rev. E **82**, 016101 (2010).

[37] A more complete explanation of a *test individual* can be found in the Supplemental Material [33] and [39].

[38] L. Decreusefond, J.-S. Dhersin, P. Moyal, and V. C. Tran, Ann. Appl. Probab. **22**, 541 (2012).

[39] J. C. Miller, Bull. Math. Biol. **74**, 2125 (2012).