Domination-time dynamics in susceptible-infected-susceptible virus competition on networks

van de Bovenkamp, R; Kuipers, FA; Van Mieghem, PFA

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Domination-time dynamics in susceptible-infected-susceptible virus competition on networks

Ruud van de Bovenkamp, Fernando Kuipers, and Piet Van Mieghem

Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, the Netherlands

When two viruses compete for healthy nodes in a simple network and both spreading rates are above the epidemic threshold, only one virus will survive. However, if we prevent the viruses from dying out, rich dynamics emerge. When both viruses are identical, one virus always dominates the other, but the dominating and dominated virus alternate. We show in the complete graph that the domination time depends on the total number of infected nodes at the beginning of the domination period and, moreover, that the distribution of the domination time decays exponentially yet slowly. When the viruses differ moderately in strength and/or speed the weaker and/or slower virus can still dominate the other but for a short time. Interestingly, depending on the number of infected nodes at the start of a domination period, being quicker can be a disadvantage.

I. INTRODUCTION

Spreading processes on networks are well studied phenomena that can be used to model the spread of diseases, rumors, opinions, and habits in populations or on online social networks, as well as the effect of marketing and product adoption [1–4]. One of the disease-spreading models that captures, for example, flulike behavior is the susceptible-infected-susceptible (SIS) model [5]. In the SIS model nodes or individuals are either healthy but susceptible, or infected and spreading the disease. The SIS model cannot be solved exactly for large networks, but much work has been done to create good approximations [6–9].

Viruses, however, do not necessarily live in isolation; a fact that has attracted attention recently. Work on virus competition either focuses on single layer networks, or on multilayer networks. Prakash *et al.* [10] proved that when two viruses compete on a single layer network, the stronger of the two will completely suppress the other, provided that the viruses are mutually immune. This is in line with the competitive exclusion principle found in ecology [11,12]. When viruses are not mutually immune, a region can exist where both viruses survive in the network [13]. Even when nodes are mutually immune, there is a region where stronger and/or quicker viruses can coexist with weaker or slower competitors in the SIR model [14]. The coexistence of SIR type epidemics is greatly influenced by node mobility in spatially structured populations [15].

In multilayer networks each virus propagates over a separate link set. Wei *et al.* [16] developed a predictor for the winner of the competition between viruses on multilayer networks and offer mitigation strategies to slow one of the two viruses down. Recently, Sahneh and Scolgio [17] improved on the understanding of virus competition in a multilayer network and showed that when the degrees in the two layers are negatively correlated, coexistence of two viruses is in fact possible, contrary to the findings in [16]. Marceau *et al.* [18] studied the effect of asymmetric an partial immunity in SIR epidemics where the spread of knowledge that a virus exists in the network helps to combat the spreading of the virus. Meyers *et al.* [19] developed a model that describes the interaction between multiple viruses (or contagions) under the assumption that a virus can also improve the spreading properties of another virus

and validated that model on messages propagating through the Twitter network.

A common feature in the study of virus spread is that viruses can become extinct. What happens, however, if one node will stick to its opinion or preference or (intentionally) stays infected? Such a model can be used to describe competing marketing campaigns where each campaign has a (small) fixed support base and explain ebbs and flows in the success of such campaigns. It can also describe the arrival of a new product or idea in a population and model its adoption starting from a fixed fraction of infected nodes. We will see in this paper that by disallowing a virus to die out, rich dynamics emerge. In line with previous research, one of the two viruses will be dominated by the other, but, contrary to what might be expected, this dominance will not last forever. Instead, viruses will alternate between being dominant and dominated, as illustrated in Fig. 1 where the fraction of infected nodes in a complete graph of 500 nodes is plotted as function of time for two viruses.

We will first describe the process in detail in Sec. II. In Sec. III we describe the modeling and simulation of the process; in the Appendix we use the GEMF model to analyze the competition process where viruses are allowed to die out. In Sec. IV we will derive the domination time of two viruses that have identical spreading properties, and finally in Sec. V we derive the domination time of two viruses that have different spreading properties.

II. MSIS PROCESS DESCRIPTION

We model the process of multiple competing viruses using the SIS epidemic spreading model. In an SIS process, a node is either healthy or infected. If a node is infected, it will spread the infection to each of its neighbors. After some time, the infected node becomes healthy again but will remain susceptible to the infection. The curing process per node is a Poisson process with rate δ and the infection process per link is a Poisson process with rate β. All processes are independent. The effective spreading rate τ is defined as the ratio between the spreading and curing rate, i.e., $\tau = \frac{\beta}{\delta}$.

The average number of infected nodes as a function of the effective spreading rate τ in the SIS process features
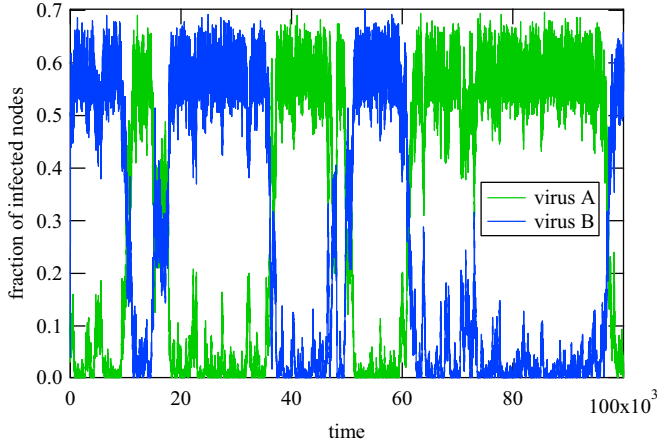
FIG. 1. (Color online) Competition between two viruses in a complete graph of 500 nodes. Each virus is prevented from dying out by reinfecting the last infected node of a virus at the moment it is cured.

two different regimes, separated by the epidemic threshold τ_c: below τ_c the virus dies out exponentially fast [20], and above τ_c the virus reaches a metastable state. In the metastable state, the average number of infected nodes hardly changes over time. The true steady state of the network is, however, the all healthy state, since the describing Markov process contains an absorbing state.

The extension from a single virus to multiple competing viruses is as follows. A node can still be either infected or healthy. The infectious state, however, is labeled with the virus type. If a node is infected with virus A, for example, it will spread the infection of type A to its healthy neighbors. The healthy state, however, is not labeled so that, irrespective of the type, nodes will only be infected with one virus at a time.

III. MODELING AND SIMULATION

We start by considering the case of two viruses in a complete graph of N nodes, K_N, and use Markov theory to completely describe the process. Although the complete graph is an extreme graph, simulations in Sec. III A show similar behavior in very different graphs. The symmetry of the complete graph allows us to use a Markov chain with a manageable number of states to describe the process.

We define N_A and N_B as the number of nodes infected with virus A and B, respectively, and N_H as the number of healthy nodes. Clearly, $N_A + N_B \leqslant N$ and $N_H = N - N_A - N_B$. A state in the Markov chain corresponds to a combination of N_A and N_B in the network. Every time the total number of infected nodes in the network changes, a state transition in the embedded Markov chain of the continuous-time Markov process occurs. The total number $\frac{(N+1)(N+2)}{2}$ of states in the Markov chain is visualized as the lower half of a rectangular grid, as illustrated in Fig. 2. The lower left corner corresponds to the all zero state; the other states in the Cartesian grid have coordinates that encode the number of infected nodes per virus type. The horizontal axis reflects the change in the number of nodes infected with virus A, and the vertical axis reflects a change in the number of nodes infected with virus B.
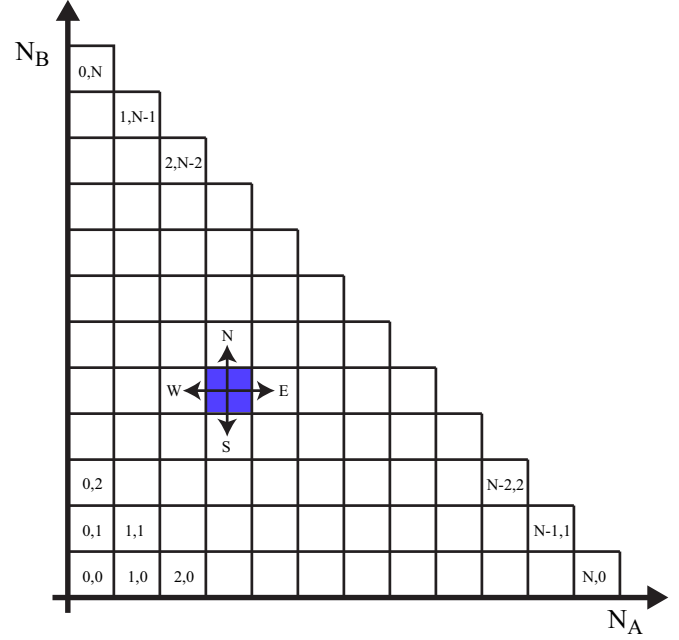

FIG. 2. (Color online) Markov chain representation of the number of infected nodes. From each state there are four different transitions possible: North, South, East, and West. The transition probabilities of a direction can be zero at the edges of the "triangle".

State changes in the Markov chain can be thought of as being either North (N_B increases), South (N_B decreases), East (N_A increases), or West (N_A decreases).

North ($\mathcal{N}(N_A,N_B)$). This state transition occurs when a virus B spreading event happens before any other event. Spreading events happen with an exponentially distributed waiting time over all $N_B N_H$ links between nodes infected with virus B and healthy nodes. The first spreading event happens after the minimum of the $N_B N_H$ waiting times, which is again an exponential random variable with rate $\beta_B N_B N_H$. The first event that is not a virus B spreading event happens after the minimum of the virus A spreading events and the virus A and B curing events, of which there are $N_A N_H$, N_A, and N_B, respectively. The waiting time until the first non-virus-B spreading event is also an exponential random variable with rate $\beta_A N_A N_H + \delta_A N_A + \delta_B N_B$. The probability that the first spreading event of virus B happens before the first nonspreading event is given by $\frac{\beta_B N_B N_H}{\beta_B N_B N_H + \beta_A N_A N_H + \delta_A N_A + \delta_B N_B}$.

South ($\mathcal{S}(N_A,N_B)$). This state transition occurs when a virus B curing event happens before any other event. Analogous to the transition to North, this probability is given by $\frac{\delta_B N_B}{\beta_B N_B N_H + \beta_A N_A N_H + \delta_A N_A + \delta_B N_B}$.

East ($\mathcal{E}(N_A,N_B)$). This state transition occurs when a virus A spreading event happens before any other event. The probability is the same as North, but with N_A and N_B interchanged: $\frac{\beta_A N_A N_H}{\beta_B N_B N_H + \beta_A N_A N_H + \delta_A N_A + \delta_B N_B}$.

West ($\mathcal{W}(N_A,N_B)$). This state transition occurs when a virus A curing event happens before any other event. The probability is the same as South, but with N_A and N_B interchanged: $\frac{\delta_A N_A}{\beta_B N_B N_H + \beta_A N_A N_H + \delta_A N_A + \delta_B N_B}$.

FIG. 3. (Color online) Structure of matrices Y_1, Y_2, and T. Gray areas consists of only zeros.

From the transition probabilities for the four directions it is clear that if N_A (or N_B) is zero, only transitions along the North-South (East-West) axis are possible. As a result, all states in Fig. 2 that are not along either the horizontal or vertical axis are transient states if we ignore the fact that the all zero state is an absorbing state. For an effective spreading rate far enough above the epidemic threshold, the probability of reaching the absorbing all zero state is very small, but the probability of reaching either of the two axes of Fig. 2 is not. This reaffirms earlier results that one virus will die out much quicker than the other.

By numbering the states appropriately, we can write the transition probability matrix P as follows [21]:

$$P = \begin{bmatrix} C_1 & 0 & 0 \\ 0 & C_2 & 0 \\ Y_1 & Y_2 & T \end{bmatrix},$$

where C_1 is a matrix containing the transition probabilities of the closed set of states with only virus A active, C_2 is a matrix containing the transition probabilities of the closed set of states with only virus B active, Y_1 and Y_2 are the matrices containing the transition probabilities from transient states to either of the two closed sets, and finally T is a $\frac{N(N-1)}{2} \times \frac{N(N-1)}{2}$ substochastic matrix containing the transition probabilities among transients states.

The structure of the submatrices of P is further detailed in Fig. 3. Matrix Y_1 is rectangular with only nonzero elements on the diagonal, with the exception of the last diagonal element. Matrix Y_2 is also rectangular but the vertical spacing is $N-1$ between the first nonzero elements and decreases for every next element. Matrix T consists of successively shrinking diagonal and off-diagonal blocks. The diagonal blocks of T are indicated in green in Fig. 3 and have the following internal structure:

$$G = \begin{bmatrix} 0 & \mathcal{E}(1,B) & & & & 0 \\ \mathcal{W}(2,B) & 0 & \mathcal{E}(2,B) & & & \\ & \mathcal{W}(3,B) & 0 & & \ddots & \\ & & \ddots & 0 & \mathcal{E}(N-B-1,B) \\ 0 & & & \mathcal{W}(N-B,B) & 0 \end{bmatrix},$$

where B indicates the block number (starting from 1 for the top left corner in T in Fig. 3). The off diagonal matrices in T are diagonal matrices:

$$R_L = \text{diag}\{\mathcal{S}(1,B+1), \mathcal{S}(2,B+1), \dots, \mathcal{S}(N-B-1,B+1)\},$$
$$R_R = \text{diag}\{\mathcal{N}(1,B), \mathcal{N}(2,B), \dots, \mathcal{N}(N-B-1,B)\},$$

where R_R is the right-hand side off-diagonal matrix; R_L is the left-hand side off-diagonal matrix.

The m-step transition probability from state i to j, that is, the probability of being in state i after m steps when starting in state j, is given by $P_{i,j}^m$. Because of the block structure of P, its powers can be expressed as

$$P^m = \begin{bmatrix} C_1^m & 0 & 0 \\ 0 & C_2^m & 0 \\ X_1 & X_2 & T^m \end{bmatrix},$$

where X_1 and X_2 are the m step transition probabilities from transient states to either of the two closed sets. Since submatrix T represents the transient states, T^m decreases with increasing m. When all elements in T^m are zero, one of the two viruses

has died out. Figure 4 shows that in a complete graph of 50 nodes the probability of both viruses being active decays exponentially after an initial period. The same results were obtained via simulations, as also indicated in Fig. 4. The simulation results are averaged over 10^7 runs. In the initial state, 20 nodes are infected with each virus. The simulation results overlap exactly with the predicted curve.

A. Perpetual competition

When both viruses are prevented from dying out, the two viruses will be in everlasting competition. In the Markov chain this can be achieved by turning the South (East) transition for all states with just one node infected with virus B (A) into the probability to stay in the same state. In simulations we reinfect a node if curing it would lead to one of the two viruses dying out.

In the simulations of the continuous-time virus competition process, the number of infected nodes with virus A and B is logged every time the total number of infected nodes changes. The moments of data gathering coincide precisely with the state changes in the Markov chain approach. For each network we simulate the competition process for 10^8 time units and log the time the process is in each state.

Figure 6 shows a heatmap representation of the steady-state distribution of the Markov chain of the virus competition in a complete graph. The number of infected nodes with virus A (B) increases along the horizontal (vertical) axis similarly as in Fig. 2. The hot spots along the axis show that the network is most likely in a state where one of the two viruses dominates the other. Either virus is equally likely to dominate the other as the symmetry along the $N_A = N_B$ axis indicates. The diagonal band connecting the two hot spots illustrates that the viruses alternate between being dominant and being dominated. The steady-state distribution confirms the behavior seen in Fig. 1. Figure 5 shows the same information as the heatmap but in such a way that multiple graph types and sizes are more easily compared. It draws the probability that the epidemic process is in a state with $(N_A - N_B)/N$ nodes infected. The probabilities $(N_A - N_B)$ are sums over all lines parallel to $N_A = N_B$ in Fig. 6.

Figure 5(a) shows the probability that the network is in a state with $(N_A - N_B)/N$ nodes infected for different spreading rates in a complete graph of 100 nodes. The different spreading rates are chosen in such a way that a single virus in this network will cause an average fraction of infected nodes between 5% and 40%. For small fractions of infected nodes both viruses will approach the absorbing state relatively frequently. The states around $N_A = N_B = 1$ are therefore visited often in the steady state and cause the high peak at $x = 0$. In the extreme case of no spreading, all states except for $N_A = N_B = 1$ are transient and the distribution features a single peak at $x = 0$. For higher spreading rates the distribution shows two peaks around the average fraction of infected nodes that characterizes the alternating domination. These two peaks are around the average fraction of infected nodes, indicating that one virus is responsible for almost all infections. They correspond to the hot spots in Fig. 6. The symmetry around zero indicates that both viruses are equally likely to dominate.

Figure 5(b) shows the distributions for complete graphs of different sizes ranging from 100 to 1600 nodes. The spreading rates are chosen in such a way that the steady-state fraction of infected nodes is 0.25. For larger networks the probability of both viruses approaching the absorbing state becomes smaller as can be seen from the lower probability of being in a state where $(N_A - N_B)/N = 0$ for larger networks. Also, the peaks in the distribution tend to be closer to the average fraction of infected nodes, in this case 0.25, for larger networks. For $N = 1600$ the peaks are located very close to the average fraction of infected nodes: around those peaks one virus is almost nonexistent, while the other has almost the entire fraction of infected nodes. The plateau between the two peaks indicates that during the crossover from one domination to the other, on average the same number of nodes are infected by either virus.

The competition behavior between two viruses is observed in various different graph types. Figure 5(c) shows the probability that the process is in a state where $(N_A - N_B)/N$ nodes are infected in five different graph types: the complete graph, the star graph, a scale-free graph grown following the preferential attachment paradigm, a connected Erdős-Rènyi random graph with link probability $\log(N)/N$, and a 32×32 square grid; all graphs, except for the grid, contain 1000 nodes. Again the effective spreading rate of both viruses is the same and on average 25% of the nodes is infected.

In the most extreme graph, the star graph, the peaks indicating one virus dominating the other are much sharper than in the complete graph. Almost no time is spent in states where both viruses have a substantial number of nodes infected. For the scale free the peaks are also markedly sharper than for the complete graph. The hub structure in both graph types probably enables a virus to dominate the other more easily and longer, as will be discussed Sec. IV. The dominated virus has to compete for the hub nodes in order to have a chance to defeat the other.

In the case of an Erdős-Rènyi random graph the difference with the complete graph is not as pronounced. The peaks are only marginally higher and the value at zero is almost exactly the same. On the contrary, the 32×32 rectangular grid shows a much higher probability to be in a state where the number of infected nodes with virus A and virus B are more or less equal. The high hop count and regular structure of the grid allows the two viruses to exist in more isolation from each other. It is still more likely to find the network in a state where one virus dominates the other, but to a much lesser extent than for the star graph.

The differences in hop count, degree distribution, and other structural properties of various graphs clearly have an influence on the competition process between the two viruses, but overall, one virus will alternately dominate the other.

IV. DOMINATION TIME OF MATCHED VIRUSES

Although the steady state of the Markov chain shows that the process will most likely be in one of two regions in the state space diagram corresponding to one virus dominating the other (see Fig. 6), it does not provide any insight in the alternating of dominating and dominated periods. In this section we investigate the domination time T of two matched viruses, i.e., $\beta_A = \beta_B, \delta_A = \delta_B$.
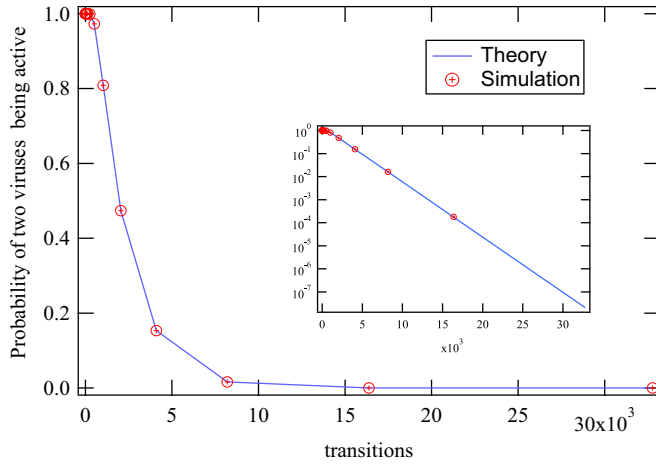
FIG. 4. (Color online) Probability that both viruses are alive as a function of the number of transitions. The solid line indicates the numerical results obtained from the transition matrix and the markers indicate the simulation results. The probability that both viruses are alive is the same as the probability that the epidemic process is in a transient state, i.e., the element wise sum of the entries of submatrix T. The inset shows the same plot on a logarithmic vertical axis and emphasizes the exponential decay of the transient state.

We say that virus A dominates virus B if there are more nodes infected with virus A than with virus B. A domination period of virus A starts when there are more nodes infected with virus A than with virus B and ends when there are more nodes infected with virus B than with virus A. The domination time is measured in the number of state changes between the start and end of a domination period. We first derive the expected domination time in Sec. IV A and compute the distribution of the domination time in Sec. IV B.

A. Expected domination time

Consider the example state space in Fig. 7 for a network of seven nodes. The Cartesian coordinates (N_A, N_B) of the state space encode the number of infected nodes with each virus. The states in which virus A dominates are blue, the states in
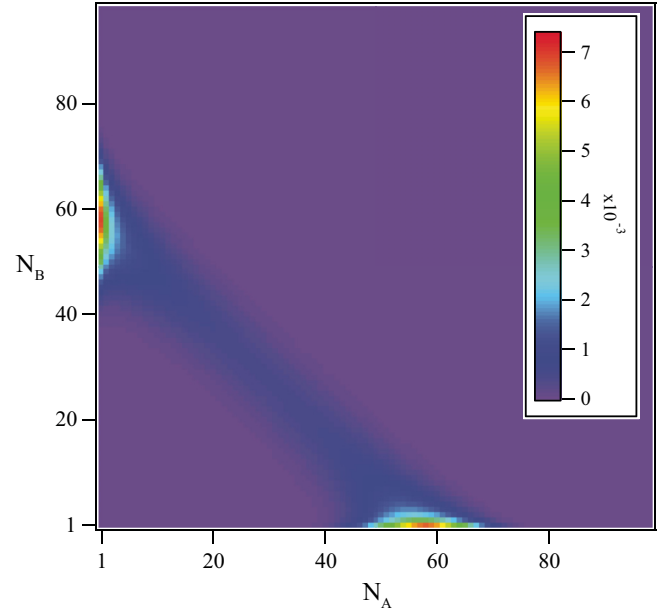
FIG. 6. (Color online) Heatmap representation of the steady-state distribution of two equally strong viruses in a complete graph of 100 nodes, and effective spreading rate $\tau = 0.025$.

which virus B dominates are green, and the states where virus A and B have the same number of infected nodes are colored yellow.

The transition probabilities in this state space can be computed using the expressions for North, South, East, and West derived in Sec. III, with the exception of the states along the axes. These states have a self-loop in the Markov transition graph to avoid the virus from dying out. The probability of staying in the same state along the axes is equal to the probability of going one more step West or South.

Let T_A be the time that virus A dominates virus B. A domination period of virus A can start at any state just above the diagonal (states 2, 8, or 12 in Fig. 7). Let's assume for now that the domination starts in state 8. The average domination time, $E[T_A]$, is the average hitting time of any state below the diagonal (the green states) given that we start in state 8. The
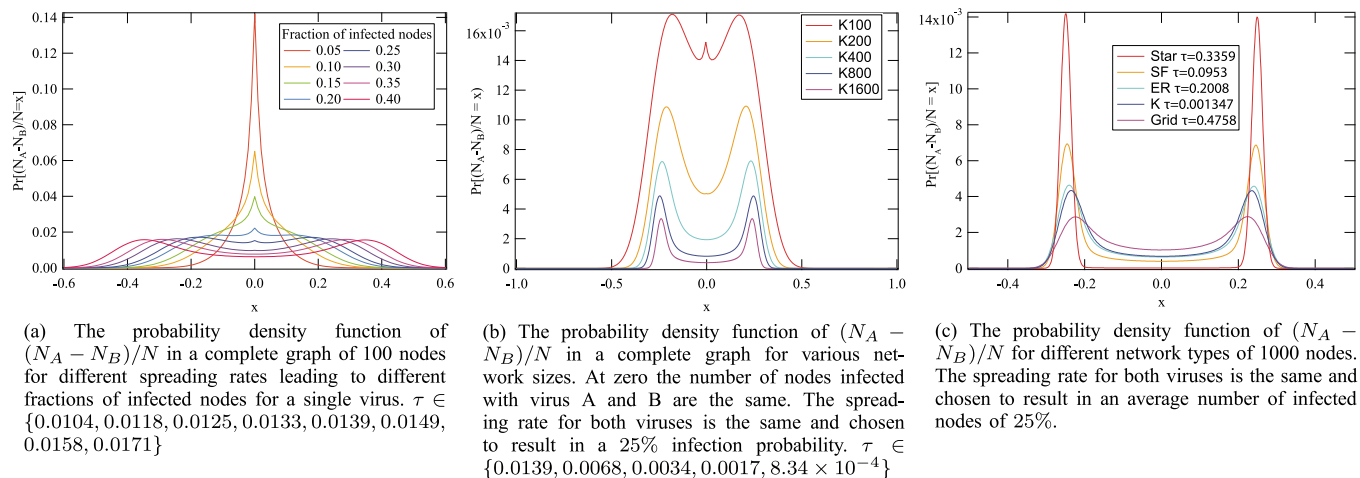
(a) The probability density function of $(N_A - N_B)/N$ in a complete graph of 100 nodes for different spreading rates leading to different fractions of infected nodes for a single virus. $\tau \in \{0.0104, 0.0118, 0.0125, 0.0133, 0.0139, 0.0149, 0.0158, 0.0171\}$

(b) The probability density function of $(N_A - N_B)/N$ in a complete graph for various network sizes. At zero the number of nodes infected with virus A and B are the same. The spreading rate for both viruses is the same and chosen to result in a 25% infection probability. $\tau \in \{0.0139, 0.0068, 0.0034, 0.0017, 8.34 \times 10^{-4}\}$

(c) The probability density function of $(N_A - N_B)/N$ for different network types of 1000 nodes. The spreading rate for both viruses is the same and chosen to result in an average number of infected nodes of 25%.

FIG. 5. (Color online) Probability density function of $(N_A - N_B)/N$ for different sizes of the complete graph (a) and different graph types of the same size (b).

average hitting time from state 8 to any green state is the value k_8 for the minimal non-negative solution [22] to the following system of linear equations:

$$k_i = 0 \quad \text{for } i \notin \{1, \dots, 12\},$$
$$k_i = 1 + \sum_{j \notin \{1, \dots, 12\}} p_{ij} k_j \quad \text{for } i \in \{1, \dots, 12\}, \tag{1}$$

where p_{ij} is the probability of going from state i to state j. The solution for k_i will generally be different for different values of i. In other words, the state in which a domination period starts influences the average duration. Figure 8 shows $E[T_A]$ in the complete graph K_{50} as a function of y_s, the number of infected nodes at the start of the domination period. Figure 8 also shows the measured average domination time of virus A as a function of y_s and the measured probability density of y_s. The measured values are averages over 10^6 domination periods. The spreading rate for both viruses was chosen in such a way that y, the expected number of infected nodes, is 33, which is about 65% of the network. Moreover, with 33 nodes infected it is possible for a domination period to start with precisely the average number of nodes infected, i.e., $y_s = y$.

Figure 8 demonstrates that our simulation results are in agreement with the numerical solution to the linear system in (1). For values of y_s that occur only rarely the simulated results are further from the numerical results because of a lack of statistics. The average domination time of a virus depends on the viral state of the network at the start of domination. When $y_s < y$ a domination period can be significantly longer than when the domination period starts close to the expected number of infected nodes. This means that in a situation with only two nodes infected, the virus that spreads first is likely to be stronger than the other for initially a rather long period. After the process converges, the influence of the initial advantage will disappear.

The effect that a domination period is longer if it starts when the number of infected nodes is smaller than the expected number of infected nodes ($y_s < y$) is more pronounced for larger spreading rates. The average T_A for periods that start with more nodes infected is only slightly larger for larger spreading rates. This means that the initial domination period for a process that starts with only two nodes infected can be very long for high spreading rates but will decrease when the other virus catches up, since it is very unlikely to

find the network in a state with few nodes infected after the start of the process.

The distributions of y_s as shown in Fig. 8 are identical to the distributions of the number of infected nodes for the three different values of the spreading rate. The distributions for y are not shown in Fig. 8 to avoid further cluttering. Because the probability of a domination period to start with y_s infected nodes is the same as the probability of the network to have y nodes infected, it is unlikely that the beginning of a domination period is related to any variations in the total number of infected nodes. If this would be the case, the number of infected nodes at the beginning of a domination period would have been higher or lower than the expected number of infected nodes. However, because of the dependency of T_A on the number of infected nodes y_s, if a domination period starts during a spell of few infected nodes, it is likely to last longer.

B. Domination-time distribution

The knowledge of the domination period is not complete without the distribution of the domination time, as the average might be a poor descriptor. We consider again the state space as illustrated in Fig. 7. To compute the distribution of the domination time, we make all the states below the diagonal (the green states) absorbing. The Markov chain that describes such a process consists of a set of transient states in which virus A dominates virus B and a set of absorbing states.

Let M be the submatrix of the transition matrix P that only contains the transitions between the transient states. In the example of Fig. 7, M is a 12×12 substochastic matrix. Let s_k be the state distribution after k steps. In our example case, $s_0 = e_8$ (an all zero vector with only the 8th element one), since we assume that the domination of virus A starts in state 8. The L_1 norm of s_k is the probability that after k steps the process is in one of the transient states. If the process is still in one of the transient states after k steps, then the domination time has to be larger than k. This means that $||s_k||_1 = \Pr[T_A > k]$, and $\Pr[T_A = k] = ||s_k||_1 - ||s_{k-1}||_1$, for $k > 0$.
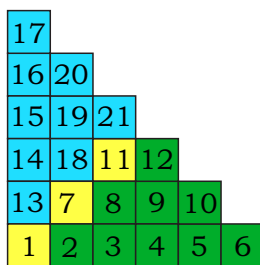

FIG. 7. (Color online) Example state space for a network with seven nodes. States where virus A dominates (the lower half) are green, states where virus B dominates (the upper half) are blue, and states where the viruses are balanced are yellow. In this state diagram state 1 corresponds to $N_A = N_B = 1$.
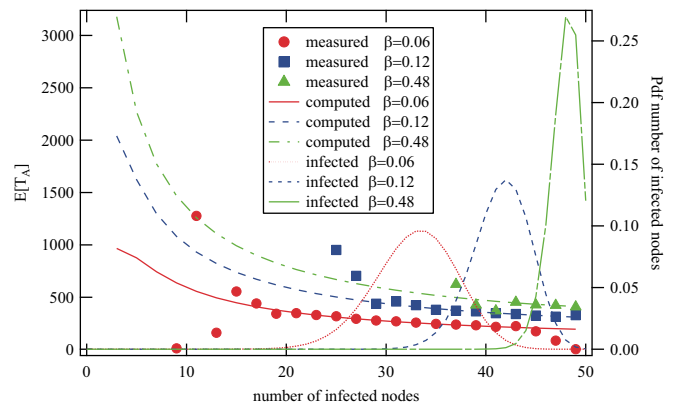

FIG. 8. (Color online) Measured and calculated domination time in a complete graph of 50 nodes as a function of the number of infected nodes at the start of the domination period for three different spreading rates β, and unit curing rate. The right-hand side vertical axis shows the measured probability density function of the number of infected nodes at the start of the domination period.

Let M be the transition matrix containing the transitions within the transient region. Assuming that all eigenvalues of M are distinct, M can be decomposed as $M = V \Lambda V^{-1}$, where V is a matrix with the eigenvectors of M as its columns and Λ is a diagonal matrix of the eigenvalues of M. The k step transition probabilities, M^k, can be expressed as $V \Lambda^k V^{-1}$. The probability of being in a transient state after k steps starting from some initial state s_0 are found [23] by summing the k step probabilities in row s_0 of M^k:

$$\Pr[T \geqslant k] = \sum_{j=1}^{N} (M^k)_{s_0, j}.$$

Because Λ is a diagonal matrix, the entries of row s_0 in $V \Lambda^k$ can be written as $v_{s_0, j} \xi_j^k$, where ξ_j is the jth eigenvalue of M. The sum over the entries of row s_0 in M^k can be written as

$$\sum_{j=1}^{N} (M^k)_{s_0, j} = \sum_{j=1}^{N} \sum_{r=1}^{N} v_{s_0, j} \xi_j^k \tilde{v}_{r, j} = \sum_{j=1}^{N} \xi_j^k c_j,$$

where $c_j = \sum_{r=1}^{N} v_{s_0, j} \tilde{v}_{r, j}$ and $\tilde{v}_{r, j}$ are the elements of V^{-1}. Using $\Pr[T = k] = \Pr[T \geqslant k] - \Pr[T \geqslant k+1]$ the probability of being in a transient state after k steps, which is the same as the domination time, can be written as

$$\Pr[T = k] = \sum_{j=1}^{N} \xi_j^k c_j (1 - \xi_j),$$

which is a weighted sum of exponentials. Because M is a substochastic matrix, all eigenvalues are smaller than 1, and the probability of being in a transient state vanishes. After sufficiently long time, only the largest eigenvalue has an influence on the distribution of the domination time leading to a purely exponential tail.

The hitting time T of the absorbing region in the embedded Markov chain is expressed in the number of transitions. In order to transform the number of state changes to time units in the continuous-time Markov process, we create a uniformized embedded Markov chain [23, p. 191]. In such a chain, self-loops ensure that transitions in the embedded chain happen at equal rates in the continuous-time process. As a result, the steady-state distributions of the uniformized embedded Markov chain and the continuous-time Markov chain are the same. The number of state changes is transformed into time by dividing by the maximum transition rate in the continuous time process.

Figure 9 shows the distribution of the domination time T (starting from $y_s = y = N/4$) in a complete graph for $N = \{50, 100, 200\}$. Small domination times are most likely, however, much longer domination times are also likely as a result of the exponential tail with small exponent. The exponential tail of all three distributions can be clearly seen on the log-lin scale. For increasing network sizes the slope of the tails becomes increasingly flat, indicating that for larger networks a virus can more easily dominate for a longer time.

The domination times for the five different graph types discussed in Sec. III A are shown in Fig. 10. The tails of all these distributions are exponential, but the differences lie in where the tail starts. For the star graph the exponential tail starts earliest. Not only are the domination periods more
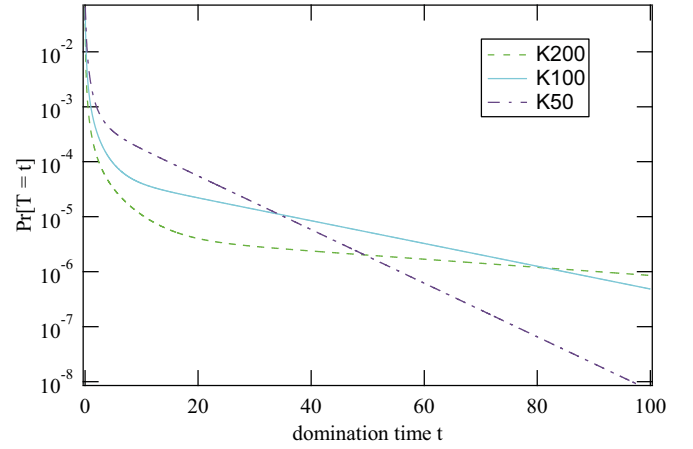

FIG. 9. (Color online) Probability distribution of the domination time T in the complete graph for three different network sizes. For each network size $y_s = y = N/4$.

pronounced for the star graph, the probability of having long domination periods is much higher than for the other graphs although the exponent is larger. For the grid the exponential tail starts the latest, indicating that small domination times are most likely. The parts of the distributions for smaller values for the domination time are very similar for all graphs except for the star graph.

V. DOMINATION TIME OF NONMATCHED VIRUSES

When the two viruses are not matched, the domination periods will no longer be the same for the two viruses. We investigate two different ways in which the two viruses can be nonmatched: in speed and in effective spreading rate. A further line of research is to investigate two viruses that have different spreading and curing rate distributions, but the same average number of spreading events during an infection period. The average number of spreading events during an infection period
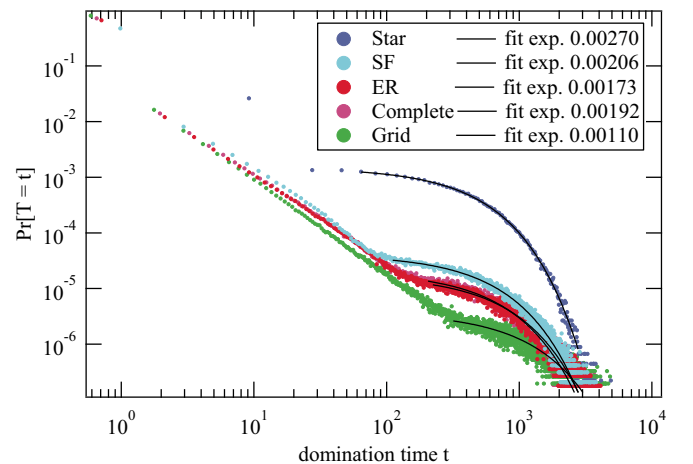

FIG. 10. (Color online) Measured probability density function of domination time T for a domination period that starts with $y_s = y$ in five different graph types of 1000 nodes. The spreading rate is chosen so that the $y = N/4$.
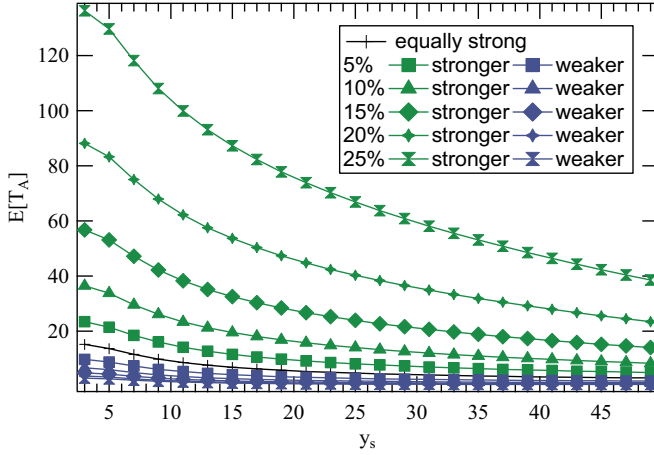
FIG. 11. (Color online) Average duration of a domination period for two viruses with different effective spreading rates as a function of y_s. The green curves indicate the stronger of the two viruses, whereas the blue curves indicate the weaker of the two.
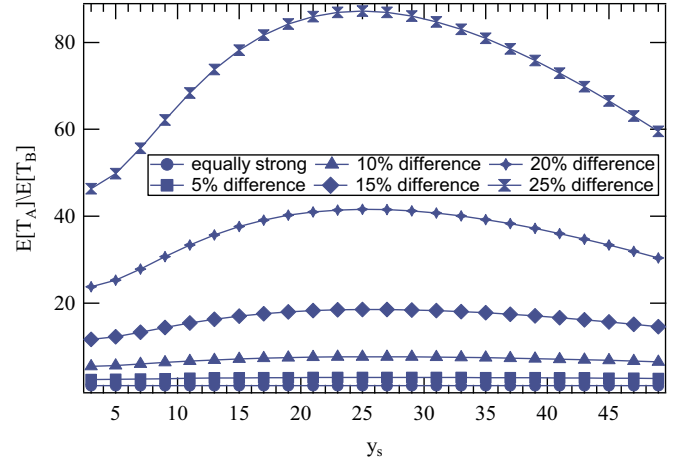

FIG. 12. (Color online) Ratio between the average duration of a domination period for the stronger virus and the duration of a domination period for the weaker virus for various differences in strength.

is a more natural way to express the spreading effectiveness in SIS processes [20,24].

A. Domination time for a stronger virus

In order to investigate the effect that one virus is stronger than the other, we vary the strength of one virus with respect to the other and compute the expected domination time. The curing and spreading rates for the two viruses are given as

$$\beta = \begin{cases} \gamma\beta_0 & \text{for virus } A, \\ \beta_0 & \text{for virus } B, \end{cases} \qquad \delta = \begin{cases} \delta_0 & \text{for virus } A, \\ \delta_0 & \text{for virus } B, \end{cases}$$

where β_0 and δ_0 are scaled for different values of γ to ensure that the average number of infected nodes stays 33.

Figure 11 shows the expected duration of a domination period for both the stronger and weaker viruses for six different strength ratios. The case in which the two viruses are matched is also shown for reference. As intuitively expected, the stronger viruses dominate the weaker virus for longer, but for values of γ not too far removed from 1, the weaker virus can still dominate the stronger one.

The effect of the difference in effective spreading rate is almost independent of y_s for small values of γ, as illustrated in Fig. 12, where the ratio $E[T_B]/E[T_A]$ is plotted for various values of γ as a function of y_s. The ratio between the expected domination time for the weaker virus and the stronger virus increases with y_s. The maximum ratio between the weaker and stronger virus occurs when y_s is smaller than the expected number of infected nodes, but not too much smaller.

Figure 13 shows the probability of the network being in a state with $(N_A - N_B)/N$ nodes infected for different values of τ_B where $\tau_B > \tau_A$. When virus B is only a few percent stronger than virus A there is still a fairly high probability of finding the network in a state where virus A dominates, but that probability drops quickly with an increasing difference in strength. Viruses need to be matched in strength within a few percent to show any real competition.

B. Domination of the quicker virus

Two viruses that have the same effective spreading rate can still differ in behavior. In this section we compare the average domination periods of viruses that differ in speed, keeping the effective spreading rate $\tau = \frac{\beta}{\delta}$ constant. The spreading and curing rates are given by

$$\beta = \begin{cases} \alpha\beta_0 & \text{for virus } A, \\ \beta_0 & \text{for virus } B, \end{cases} \qquad \delta = \begin{cases} \alpha\delta_0 & \text{for virus } A, \\ \delta_0 & \text{for virus } B, \end{cases}$$

where β_0 and δ_0 are the starting values for the virus process such that the average number of infected nodes is 33. Contrary to the case of the varying mutual strength, the difference in speed does not lead to a difference in the *total* number of infected nodes.

Figure 14 plots the average domination time as a function of y_s for various values of α. As expected, the quicker virus has an advantage for almost all values of y_s, but surprisingly, when $y_s > y$, being quicker becomes disadvantageous. This
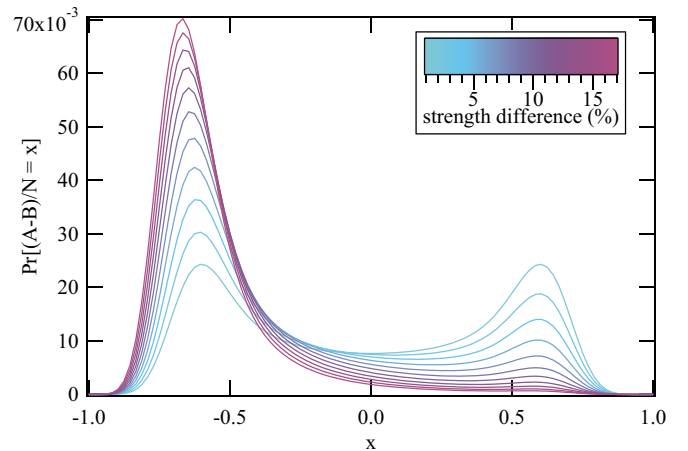

FIG. 13. (Color online) Probability density function of $(N_A - N_B)/N$ for a complete graph of 50 nodes where $\tau_B = \gamma\tau_A$, with $1 \leqslant \gamma \leqslant 1.17$.
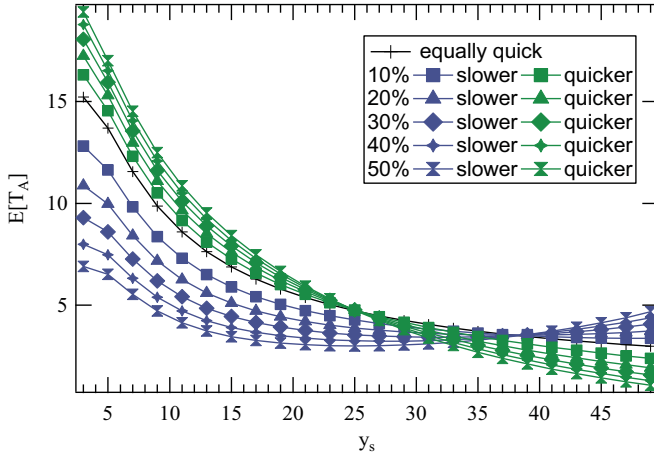
FIG. 14. (Color online) Average duration of a domination period for two viruses with different spreading speed as a function of the number of infected nodes at the start of the domination period. The green curves indicate the quicker of the two viruses, whereas the blue curves indicate the slower of the two.

can be explained by the fact that when a domination period starts during a period when the number of infected nodes is above the average there is a driving force that will push the number of infected nodes back to the mean. Since the quicker virus dominates, there are more nodes infected with the quicker virus. All these nodes have a smaller curing time and it is therefore likely that the number of nodes infected with the quicker virus will decrease before the nodes infected with the slower virus will decrease thereby shortening the expected domination time of the stronger virus.

Because of the disadvantage of being a quicker virus for values of y_s that are above the expected number of infected nodes, the effect of α on the ratio between the expected domination time of a slower virus and a quicker virus is not constant over y_s as for viruses that differ moderately in strength. Figure 15 shows the ratio $E[T_B]/E[T_A]$ for different
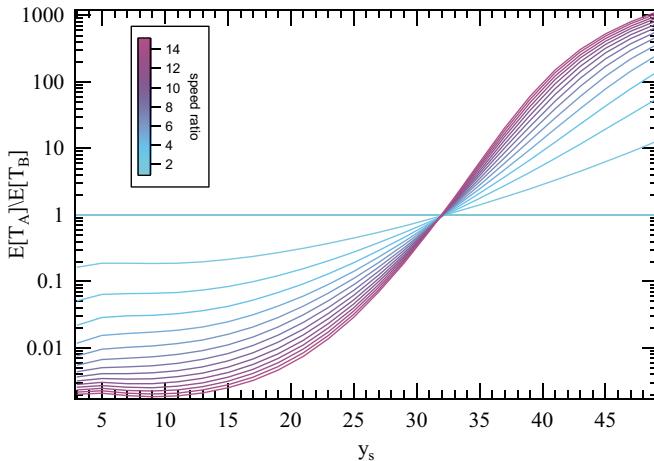

FIG. 15. (Color online) Ratio between the average duration of a domination period for the slower virus and the duration of a domination period for the quicker virus for various differences in speed: $\beta_B = \alpha\beta_A$ for $1 \leqslant \alpha \leqslant 15$.

values of α as a function of y_s. The shapes of these curves are markedly different from those of the strength difference in Fig. 12. This means that although it is possible to balance a stronger virus with a quicker virus for a *single* value of y_s, it is not possible for the entire range of y_s. Also, the curves are symmetric around $y_s = y$, resulting in a balance in the total time the quicker virus dominates the slower and vice versa.

VI. CONCLUSION

When two viruses compete for healthy nodes in a network, we have shown that one of the two viruses dies out exponentially fast even when both viruses have an effective spreading rate above the epidemic threshold. However, when we prevent two equally strong viruses from dying out, a rich dynamic process emerges where one of the two viruses dominates the other but loses that dominance after some time. We have shown that the average domination time of a virus depends on the number of infected nodes at the beginning of the domination period and that it can be computed by solving a linear system. The distribution of the domination time can also be computed numerically.

When the two viruses are not balanced but differ in either strength or speed, the domination times will also be unbalanced. If the differences are not too big, the weaker and/or slower viruses will still periodically dominate their quicker and/or stronger rivals, but only for a short time. Because the effect of speed and strength as a function of the number of infected nodes at the start of the domination time differs, it is impossible to balance a stronger virus with a quicker rival for all values of y_s, but it is possible to balance them for a single value of y_s. Contrary to being stronger than a rival, being quicker is not always a benefit but depends on both how much quicker the virus is and how many nodes are infected at the beginning of a domination period. Interesting future work includes the effect of non-Markovian spreading properties of both viruses. In this case the strength of two viruses can be matched, while they differ in distributions for the interarrival rate of spreading and curing events.

APPENDIX: GENERALIZED EPIDEMIC MEAN-FIELD MODEL

Recently, the generalized epidemic mean-field model (GEMF) has been introduced by Darabi Sahneh *et al.* [25] as a generalization of NIMFA, the N-intertwined mean-field approximation [25,26]. GEMF uses a node level description of the spreading process to arrive, through a mean-field approximation, at a set of nonlinear ordinary differential equations that describe the network state as a whole. Nodes can be in one of several states (or compartments) and interact through a multilayer network depending on their state. The node level description of the process is given by a set of
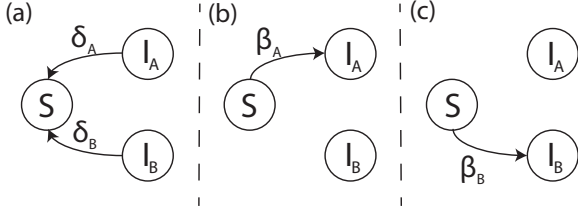
FIG. 16. Transition rate graph of the MSIS process. (a) Nodal transition rate graph: the three states are susceptible (S), infected with virus A (I_A), and infected with virus B (I_B). A node moves from state I_A to S with curing rate δ_A and from I_B to S with curing rate δ_B. (b) Edge-based transition graph for type A: a node moves from state S to state I_A with spreading rate β_A for each incident link from a node in I_A. (c) Edge-based transition graph for type B: a node moves from state S to I_B with spreading rate β_B for each incident link from a node in I_B. The contact networks for I_A and I_B are identical.

transition rate graphs; each transition rate graph can either be a node-based rate graph, if the rate is independent of the state of a node's neighbors, or an edge-based transition graph if the transition does depend on the state of a node's neighbors. In the case of the MSIS process, curing is a node-based transition, whereas spreading is an edge-based transition. The edge-based transitions are dependent on an influencer state. The influencer state is a label that indicates that node i's transition from one state to another depends on the number of neighbors of node i in the indicated state. The MSIS process with two viruses has three states in the GEMF model; the susceptible state (S), the infected with virus A state (I_A), and the infected with virus B state (I_B). Figure 16 shows the transition rate graphs for this MSIS process as follows. (a) The nodal rate graph representing the curing of infected nodes, which happens with curing rate δ_A for virus A and curing rate δ_B for virus B. (b) The transition from state S to I_A, which happens with rate β_A per link and has influencer state I_A. (c) The transition from state S to I_B, which happens with rate β_B per link and has influencer state I_B. The interaction networks for states I_A and I_B are identical.

The state $x_i(t)$ of node i at time t is represented by the standard unit vector corresponding to the state i. The expected value of $x_i(t)$ is defined as [25, Eq. (2)]

$$E[x_i(t)] = [\Pr[x_i(t) = e_i], \dots, \Pr[x_i(t) = e_M]]^T \triangleq v_i(t),$$

where e_i is the ith standard unit vector and M the number of states. The governing equations of the GEMF model follow from the transition graphs [25, Eq. (27)]:

$$\frac{dv_i}{dt} = -Q_\delta^T v_i - \sum_{l=1}^{L} \left(\sum_{j=1}^{N} a_{i,j|l} v_{j,q_l} \right) Q_{\beta_l}^T v_i, i = \{1, \dots, N\},$$

$$(A1)$$

where Q_δ is the Laplacian matrix of the nodal transition rate graph, Q_{β_l} are the edge transition graphs for the various layers, and $a_{i,j|l}$ are the adjacency matrix elements of the network describing the interactions on layer l, q_l is the influencer state at layer l, L is the number of network layers, and N is the number of nodes in the network.

For the two-virus process described by the transition graphs in Fig. 16, we have the following Laplacian matrices:

$$Q_\delta = \begin{bmatrix} 0 & 0 & 0 \\ -\delta_A & \delta_A & 0 \\ -\delta_B & 0 & \delta_B \end{bmatrix}, \quad Q_{\beta_A} = \begin{bmatrix} \beta_A & -\beta_A & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$Q_{\beta_B} = \begin{bmatrix} \beta_B & 0 & -\beta_B \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (A2)$$

The interaction networks for the two infectious states, I_A and I_B, are identical for the 2-MSIS process, so we can drop the index l from $a_{i,j|l}$. As indicated above, the influencer states for I_A and I_B are the states themselves. Equation (A1) in the case of 2-MSIS reduces to

$$\dot{v}_i = -Q_\delta^T v_i - \sum_{j=1}^{N} a_{ij} v_{j,I_A} Q_{\beta_A}^T v_i - \sum_{j=1}^{N} a_{ij} v_{j,I_B} Q_{\beta_B}^T v_i,$$

$$(A3)$$

which we can expand to

$$\begin{bmatrix} \dot{S_i} \\ I_{A_i} \\ I_{B_i} \end{bmatrix} = - \begin{bmatrix} 0 & -\delta_A & -\delta_B \\ 0 & \delta_A & 0 \\ 0 & 0 & \delta_B \end{bmatrix} \begin{bmatrix} S_i \\ I_A \\ I_B \end{bmatrix}$$

$$- \sum_{j=1}^{N} a_{ij} I_{A_j} \begin{bmatrix} \beta_A & 0 & 0 \\ -\beta_A & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} S_i \\ I_{A_i} \\ I_{B_i} \end{bmatrix}$$

$$- \sum_{j=1}^{N} a_{ij} I_{B_j} \begin{bmatrix} \beta_B & 0 & 0 \\ 0 & 0 & 0 \\ -\beta_B & 0 & 0 \end{bmatrix} \begin{bmatrix} S_i \\ I_{A_i} \\ I_{B_i} \end{bmatrix}. \quad (A4)$$

If $\beta_A = \beta_B$ and $\delta_A = \delta_B$, the total number of infected nodes equals the number of infected nodes of a single SIS process. We take the first row in Eq. (A4):

$$\frac{dS_i}{dt} = \delta_A I_{A_i} + \delta_B I_{B_i} - \sum_{j=1}^{N} a_{ij} I_{A_j} \beta_A S_i - \sum_{j=1}^{N} a_{ij} I_{B_j} \beta_B S_i$$

$$= \delta(I_{A_i} + I_{B_i}) - S_i \beta \sum_{j=1}^{N} a_{ij} \left(I_{A_j} + I_{B_j} \right). \quad (A5)$$

After using $I = I_A + I_B$ and $S + I = 1$, Eq. (A5) reduces to

$$\frac{dS_i}{dt} = \delta(1 - S_i) - S_i \beta \sum_{j=1}^{N} a_{ij} I_j, \quad (A6)$$

which is the governing equation of the N-intertwined mean-field approximation of the normal SIS processes; see, for example, [26, Eq. (1)].

1. Evaluation of the GEMF model

To verify the correctness of the GEMF model description of the MSIS process, we compare the numerical solution of Eq. (A3) to simulations in a complete graph of 500 nodes. The average number of infected nodes per virus type as a function of time can be computed from Eq. (A3) as the average number of nodes in one of the infectious states. For the simulations, we
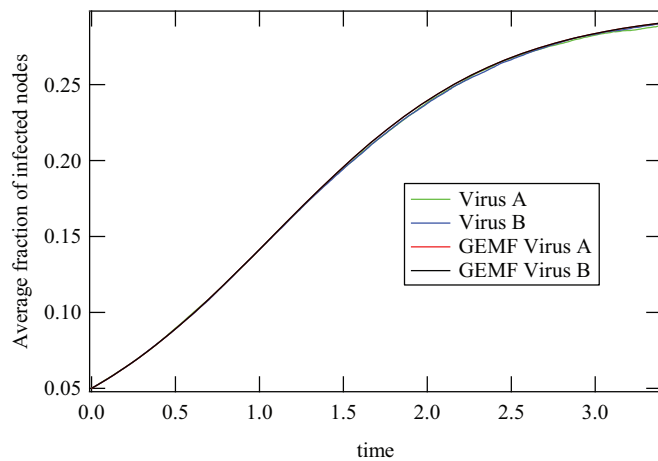
FIG. 17. (Color online) Comparison between the average number of infected nodes in the GEMF model simulations in a complete graph of 500 nodes. The simulation results are averaged over 100 000 runs.

have averaged the number of infected nodes at regular sample intervals over 100 000 runs using the event based simulator

first described in [27]. At the start of each simulation 20% of the nodes are infected with virus *A* and another 20% of the nodes are infected with virus *B*.

In Fig. 17 the average fraction of infected as computed from Eq. (A3) and simulations are plotted as a function of time. The GEMF model and the simulations both coincide and show that the average fraction of infected nodes is the same for viruses *A* and *B*. Yet, Fig. 17 should be interpreted with care. Just as mean-field models for normal SIS, the GEMF model for MSIS cannot describe the process of a virus dying out. In reality, one of the two viruses rapidly dies out as a result of the competition.

The effects of the competition can be highlighted by preventing the viruses from dying out. This is done by reinfecting the last node of a virus type at the moment it is cured. Figure 1 shows the fraction of infected nodes per virus type as a function of time, again in a complete graph of 500 nodes. In contrast to Fig. 17, the fraction of infected nodes is doubled for the dominant virus, while it is very small for the dominated virus. Interestingly, the two viruses alternate from being dominant to being dominated. This behavior is further discussed in Sec. III A.

[1] C. Doerr, N. Blenn, and P. Van Mieghem, PLoS ONE **8**, 1 (2013).

[2] J. Leskovec, L. A. Adamic, and B. A. Huberman, ACM Trans. Web **1**, 5 (2007).

[3] D. Gruhl, R. Guha, D. Liben-Nowell, and A. Tomkins, in *Proceedings of the 13th International Conference on World Wide Web*, ser. WWW '04, New York, NY (ACM, New York, 2004), pp. 491–501.

[4] J. Leskovec, M. McGlohon, C. Faloutsos, N. Glance, and M. Hurst, in *Proceedings of the 2007 SIAM International Conference on Data Mining, 2007* (SIAM, Philadelphia, 2007), pp. 551–556.

[5] R. Anderson and R. May, *Infectious Diseases of Humans: Dynamics and Control* (Oxford University Press, New York, 1991).

[6] P. Van Mieghem, J. Omic, and R. Kooij, IEEE/ACM Trans. Netw. **17**, 1 (2009).

[7] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos, in *Proceedings of the 22nd International Symposium on Reliable Distributed Systems, 2003* (IEEE, Piscataway, NJ, 2003), pp. 25–34.

[8] A. Ganesh, L. Massoulie, and D. Towsley, in *Proceedings of the IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM 2005* (IEEE, New York, 2005), Vol. 2, pp. 1455–1466.

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

[10] B. A. Prakash, A. Beutel, R. Rosenfeld, and C. Faloutsos, in *Proceedings of the 21st International Conference on World Wide Web*, ser. WWW '12, New York, NY (ACM, New York, 2012), pp. 1037–1046.

[11] C. Castillo-Chavez, W. Huang, and J. Li, SIAM J. Appl. Math. **59**, 1790 (1999).

[12] A. Ackleh and L. Allen, Discr. Contin. Dynam. Syst. Ser. B **5**, 175 (2005).

[13] A. Beutel, B. A. Prakash, R. Rosenfeld, and C. Faloutsos, in *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '12, New York, NY (ACM, New York, 2012), pp. 426–434.

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

[15] C. Poletto, S. Meloni, V. Colizza, Y. Moreno, and A. Vespignani, PLoS Comput. Biol. **9**, e1003169 (2013).

[16] X. Wei, N. Valler, B. Prakash, I. Neamtiu, M. Faloutsos, and C. Faloutsos, IEEE J. Sel. Areas Commun. **31**, 1049 (2013).

[17] F. Shahneh and C. Scoglio, arXiv:1308.4880.

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

[19] S. A. Myers and J. Leskovec, Data Mining **12**, 539 (2012).

[20] P. Van Mieghem and R. van de Bovenkamp, Phys. Rev. Lett. **110**, 108701 (2013).

[21] F. Gebali, *Analysis of Computer and Communication Networks* (Springer, New York, 2008).

[22] J. Norris, *Markov Chains*, Cambridge Series on Statistical and Probabilistic Mathematics (Cambridge University Press, Cambridge, UK, 2005).

[23] P. Van Mieghem, *Performance Analysis for Communications Networks and Systems* (Cambridge University Press, Cambridge, UK, 2006).

[24] E. Cator, R. van de Bovenkamp, and P. Van Mieghem, Phys. Rev. E **87**, 062816 (2013).

[25] F. Darabi Sahneh, C. Scoglio, and P. Van Mieghem, IEEE/ACM Trans. Netw. **21**, 1609 (2013).

[26] P. Van Mieghem, Computing **93**, 147 (2011).

[27] C. Li, R. van de Bovenkamp, and P. Van Mieghem, Phys. Rev. E **86**, 026116 (2012).