

**Survival time of the susceptible-infected-susceptible infection process on a graph**

Ruud van de Bovenkamp and Piet Van Mieghem

*Delft University of Technology, 2628 CD Delft, The Netherlands*

(Received 22 July 2014; published 15 September 2015)

The survival time  $T$  is the longest time that a virus, a meme, or a failure can propagate in a network. Using the hitting time of the absorbing state in an uniformized embedded Markov chain of the continuous-time susceptible-infected-susceptible (SIS) Markov process, we derive an exact expression for the average survival time  $E[T]$  of a virus in the complete graph  $K_N$  and the star graph  $K_{1,N-1}$ . By using the survival time, instead of the average fraction of infected nodes, we propose a new method to approximate the SIS epidemic threshold  $\tau_c$  that, at least for  $K_N$  and  $K_{1,N-1}$ , correctly scales with the number of nodes  $N$  and that is superior to the epidemic threshold  $\tau_c^{(1)} = \frac{1}{\lambda_1}$  of the N-intertwined mean-field approximation, where  $\lambda_1$  is the spectral radius of the adjacency matrix of the graph  $G$ . Although this new approximation of the epidemic threshold offers a more intuitive understanding of the SIS process, it remains difficult to compare outbreaks in different graph types. For example, the survival in an arbitrary graph seems upper bounded by the complete graph and lower bounded by the star graph as a function of the normalized effective infection rate  $\frac{\tau}{\tau_c}$ . However, when the average fraction of infected nodes is used as a basis for comparison, the virus will survive in the star graph longer than in any other graph, making the star graph the worst-case graph instead of the complete graph. Finally, in non-Markovian SIS, the distribution of the spreading attempts over the infectious period of a node influences the survival time, even if the expected number of spreading attempts during an infectious period (the non-Markovian equivalent of the effective infection rate) is kept constant. Both early and late infection attempts lead to shorter survival times. Interestingly, just as in Markovian SIS, the survival times appear to be exponentially distributed, regardless of the infection and curing time distributions.

DOI: [10.1103/PhysRevE.92.032806](https://doi.org/10.1103/PhysRevE.92.032806)

PACS number(s): 64.60.aq, 05.40.-a, 87.19.X-

**I. INTRODUCTION**

Spreading phenomena on networks, such as social contagion; the spread of messages, memes, and tweets on online social networks; the propagation of failures in infrastructural networks; as well as biological diseases, play a paramount role and can impact our society seriously. Unfortunately, their mechanisms are currently far from sufficiently understood.

Most studies on the classical epidemiological models such as susceptible-infected-susceptible (SIS) and susceptible-infected-recovered (SIR) [1,2], which may give a first-order insight into the emergent behavior on networks, have predominantly focused on the determination of the number or fraction of infected nodes after a long time and on the epidemic threshold [3–5]. The epidemic threshold, the characterizer of a phase transition, separates the regime where an infection dies out quickly from the regime where the average number of infected nodes is stable over a long time, called the metastable or quasistationary state.

A first-order mean-field approximation of the epidemic threshold  $\tau_c^{(1)} = \frac{1}{\lambda_1(A)}$ , where  $\lambda_1(A)$  is the largest eigenvalue of the adjacency matrix  $A$ , was first proposed by Wang *et al.* [6] and rigorously proved by Van Mieghem *et al.* in Refs. [5,7] and later appeared in the physics community [8]. Van Mieghem *et al.* [5] also showed that this mean-field threshold lower bounds the “in reality observed” epidemic threshold,  $\tau_c^{(1)} = \frac{1}{\lambda_1(A)} \leq \tau_c$ . A more accurate lower bound  $\tau_c \geq \tau_c^{(2)} \geq \tau_c^{(1)}$ , where  $\tau_c^{(2)}$  is the second-order mean-field threshold, was derived in Ref. [9]. Alternatively, in the work of Pastor-Satorras *et al.* [3,10], a heterogeneous mean-field approximation of the epidemic threshold was derived as  $\tau_c^{\text{HMF}} = E[D]/E[D^2]$ , where  $D$  is the degree of a randomly chosen node in  $G$ . For power-law degree graphs, there

is evidence of multiple phase transitions [11] in the SIS process.

Here we deepen our understanding by considering the time aspect of an epidemic. Our main concern is the following question: How long will the virus stay in this metastable state? Below the epidemic threshold, the infectious process dies out exponentially fast with time [12,13]. However, around and above the epidemic threshold, relatively little is known about the time dynamics of epidemics, even of the relatively simple, Markovian SIS model on networks: What is the role of the network topology on the average time that an epidemic lasts? How does this time vary with the size of the network? What is the effect of the initial number of infected nodes?

Before specifying our new contributions, we define the SIS epidemic process on a graph  $G(N, L)$  with  $N$  nodes and  $L$  links. In the SIS process, a node can be in one of two states: infected or healthy. Infected nodes spread the infection to their healthy neighbors with the infection rate  $\beta$  per link and they cure with the rate  $\delta$  back to healthy but susceptible nodes. In the Markovian SIS processes, both the infection and curing processes are independent Poisson processes, and the ratio between the infection and curing rate is defined as the effective infection rate  $\tau = \frac{\beta}{\delta}$ , also called the spreading rate [2].

In Sec. II, based on the hitting time of the absorbing state in the uniformized embedded Markov chain, we derive an exact solution for the average survival time in the complete graph (Sec. II A) and the star graph (Sec. II B). For the complete graph, we show how our results lead to new interpretations of previous work. In addition, we propose a new, precise method to approximate the epidemic threshold (Sec. II C), based on the formalism of the Markovian hitting time and evaluate the accuracy of the method on the known, exact results of the epidemic threshold for the complete graph and the star [14].

These analytic results provide the major insight for simulation results on other types of graphs, presented in Sec. III.

Finally, in Sec. IV, we extend the Markovian approach towards more real-world phenomena, that are likely not Markovian [15,16]. We show that in non-Markovian spreading, a heavy-tailed distribution for the infection rate leads to shorter survival times but the survival time distribution still has an exponential tail. Conversely, for rate distributions with a steeper tail, the expected survival time increases. As a function of the shape parameter  $\alpha$ , which controls how quickly the Weibull distribution (8) decays, the average survival time peaks just above 1, that is, just after the exponential case.

## II. SURVIVAL TIME VIA HITTING TIME

The classical SIS process on a graph is completely described by a  $2^N$  state continuous-time Markov chain [5,17], where  $N$  is the number of nodes in the network. Each state in the SIS Markov chain represents a particular configuration of infected nodes. Since a node is either infected or healthy, each state can be expressed as a binary string of length  $N$  in which bit  $i$  corresponds to the state of node  $i$ , leading to a total of  $2^N$  states. Only state transitions that differ in one bit are possible. The state in which the virus has died out, or the string contains only zeros, is an absorbing state in the Markov chain. The average survival time of the SIS process can be computed via the average hitting time of the absorbing state. Computing the hitting time for a Markov chain with such a large state space is only feasible for very small graphs. For graphs where the infectious state shows some form of symmetry, however, the number of states can be dramatically reduced. Unfortunately, symmetry in the graph structure or node degree alone is not enough to reduce the size of the state space; symmetry in the infectious state of the network is required. In the complete graph  $K_N$ , for example, every infected node links to all healthy nodes. As a result, every state with  $n$  infected nodes has transitions to all states with  $n - 1$ , as well as to all states with  $n + 1$ , infected nodes, leading to a birth and death process with  $N$  states [12]. In this section, we exploit the symmetry in the infectious state of the complete graph and the star graph, which possess a simplified Markov chain with  $O(N)$  states, to compute the hitting time of the absorbing state. Clearly, the hitting time of the absorbing state depends on the starting state. The worst-case average survival time occurs when initially all nodes are infected. This paper mainly focuses on the worst-case survival time. Using the hitting time, we derive an alternative expression for the epidemic threshold, which we compare to the exact results [14] on the complete and star graph.

We use an embedded Markov chain to transform the continuous-time Markov chain to a discrete-time one. The embedded Markov chain of a continuous-time process contains the transition probabilities at the time of a transition, but no longer contains the precise timing of the events (see Ref. [12]). The average hitting time of the absorbing state in the embedded chain, starting from the all-infected state, gives the average number of transitions between the initial state and reaching the absorbing state. However, the time between transitions is unknown. By introducing self-loops in the embedded Markov chain, the average transition rate from state  $i$  to  $j$  ( $i \neq j$ ) in the

embedded chain is made identical to the transition rate in the original continuous-time Markov chain. The transition rate, including self-loops, is uniform, which enables us to relate the number of transitions to time.

The transition matrix of the uniformized embedded Markov chain in units of  $\phi$  of a continuous-time Markov chain is given by  $S(\phi) = J + \frac{Q}{\phi}$ , where  $Q = [q_{ij}]$  is the infinitesimal generator of the continuous-time Markov chain [12] and  $J$  is the identity matrix. The off-diagonal elements  $q_{ij}$  of the infinitesimal generator contain the transition rates from state  $i$  to state  $j$ , while the diagonal elements are given by  $-\sum_{j=1, j \neq i} q_{ij}$ . Transitions in the uniformized Markov chain all occur with the same rate  $\phi \geq \max_i q_i$ , where  $q_i$  is the total outgoing rate of state  $i$ :  $q_i = \sum_{j=1, j \neq i} q_{ij}$ , with  $q_{ij}$  the transition rate from state  $i$  to state  $j$ .

The average hitting time of the absorbing state in a discrete Markov chain is given by the minimal non-negative solution of the following system [12,18]:

$$\begin{aligned} w_1 &= 0 \\ w_i &= 1 + \sum s_{ij}(\phi)w_j \quad \text{for } i \neq 1, \end{aligned} \quad (1)$$

where  $w_i$  is the average hitting time of the absorbing state starting from state  $i$  and  $s_{ij}$  is the transition probability from state  $i$  to  $j$  in the uniformized embedded Markov chain.

The system of equations (1) can be written as  $Kw = b$ , where  $K = J - S(\phi)$ . The vector  $b$  is defined as  $b = \mathbf{u} - e_1$ , where  $\mathbf{u}$  is the all-one vector and  $e_1$  is the first standard basis vector. The minimal non-negative solution of system (1) for  $w_i$  will give the average hitting time  $E[T]$  of the absorbing state when starting in state  $i$  in units of  $\phi^{-1}$ , whereas multiplying  $K$  by  $\phi$  will give the average hitting time in units of 1. Therefore, we solve the system

$$\phi K w = b, \quad (2)$$

where  $\phi K$  simplifies to  $\phi K = \phi(J - J + \frac{Q}{\phi}) = -Q$ .

### A. The complete graph

The Markov transition graph of the SIS Markov chain on the complete graph  $K_N$  is shown in Fig. 16. Solving system (2) for the complete graph (see Appendixes A and C) yields the average survival time

$$E[T] = \sum_{j=1}^N \frac{x_j}{j}, \quad (3)$$

where  $x_j$  obeys the recursive relation

$$x_{j+1} = x_j(N - j)\tau + 1, \quad (4)$$

with initial conditions  $x_1 = 1$ ,  $x_j = 0$  for  $j < 0$ .

The result (3) is connected in a surprising way to earlier results. Given that a continuous-time Markov process, containing an absorbing state, starts from the quasistationary distribution, Artalejo [19] has demonstrated that its survival time  $T$  is exponentially distributed as  $\Pr[T \leq t] = 1 - \exp(-t\zeta)$ , where  $\zeta$  is the second largest eigenvalue of the infinitesimal generator  $Q$  of the Markov chain. Fill [20] and Miclo [21] show that, for an irreducible, continuous-time birth-and-death process with  $N$  states (such as the SIS process on the complete

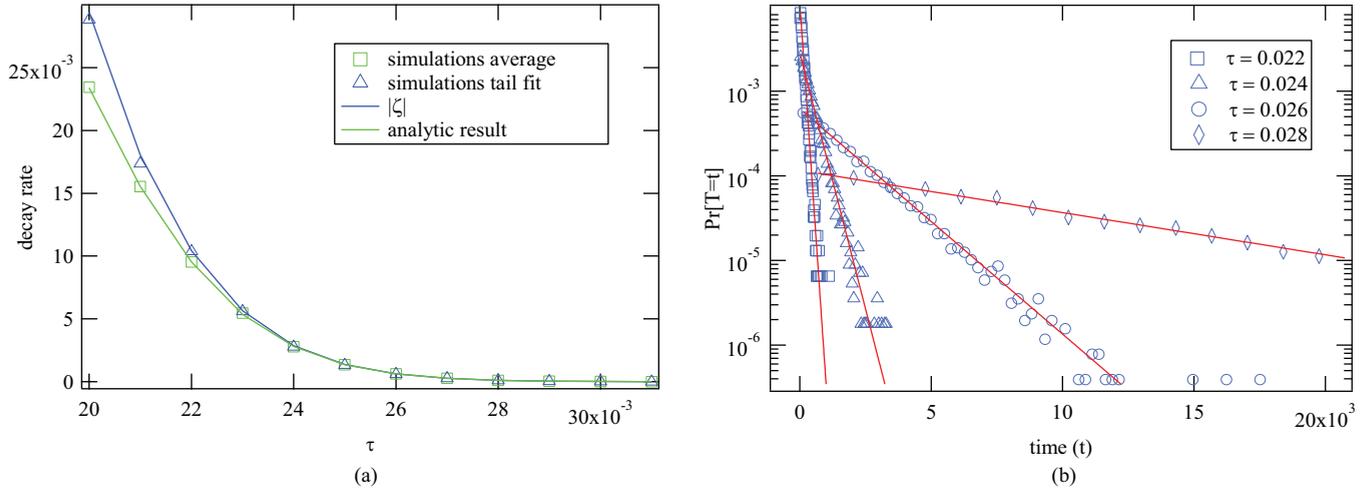


FIG. 1. (Color online) Decay rate in a complete graph of 64 nodes (a) and histograms of the survival time (b). (a) The decay rate  $|\zeta|$  as a function of the effective infection rate  $\tau$  in a complete graph of 64 nodes determined in four different ways: average survival time (squares), fitting the histogram of the survival time (triangles),  $\zeta$  (green line), and Eq. (3) (blue line). (b) Histograms of the survival time in units of  $\delta = 1$  in a complete graph of 64 nodes for effective infection rate  $\tau$  ranging from 0.022 to 0.028. The epidemic threshold in  $K_N$  is close to  $\frac{1}{N-1} = 0.016$ .

graph  $K_N$ , the hitting time of the absorbing state is distributed as the sum of  $N$  independent exponential random variables with parameters equal to the eigenvalues of  $Q$ . Since the spectrum of  $Q$  contains information about the survival time of the SIS process, The author of Ref. [22] recently used a Lagrange series to derive the accurate expression, for large  $N$ , of the decay rate  $\zeta < 0$  of the survival time for effective infection rates  $\tau > \tau_c$  in  $K_N$ . In particular, the average survival time  $E[T]$  of the SIS process in  $K_N$  is shown [22] to be approximately equal to  $|\zeta^{-1}|$ , with

$$-\zeta = \frac{\delta}{F(\tau)} + O\left(\frac{N^2 \log N}{(\tau N)^{2N-1}}\right), \quad (5)$$

where

$$F(\tau) = \sum_{j=1}^N \sum_{r=0}^{j-1} \frac{(N-j+r)!}{j(N-j)!} \tau^r. \quad (6)$$

Remarkably, as proved in Appendix C, for a Markovian SIS epidemic process on the complete graph  $K_N$ , the average survival time  $E[T]$  is *exactly* equal to  $F(\tau)$  for all effective infection rates  $\tau$ , and thus not only for  $\tau > \tau_c$ .

Equation (6) appears in the first term in the Lagrange series (5) of the second largest eigenvalue  $\zeta$  of the infinitesimal generator  $Q$  of the continuous-time SIS Markov process on  $K_N$  (discussed further in Sec. II A) and is expected to be close to  $\zeta$ . In fact, for an effective infection rate  $\tau > \tau_c$ , the order term in (5) tends *exponentially* fast in  $N$  to zero!

To show the difference between the various ways of determining the survival time of the SIS process in  $K_N$ , we determine the decay rate of the survival time in a complete graph of 64 nodes in four different ways: (i, blue line) the second largest eigenvalue  $\zeta$  of the infinitesimal generator  $Q$ , (ii, triangular markers) an exponential fit to the histogram of simulated survival times, (iii, square markers) the reciprocal of the average of the simulated survival times, (iv, green line) computed from (3) that is numerically superior to the double

sum in (6) for large  $N$ . All simulations start in the all-infected state and the time till extinction is measured.

Figure 1(a) shows that the decay rate found by the four different methods all converge to the same value as the effective infection rate  $\tau$  increases. For relatively small values of  $\tau$ , however, the four methods split into two groups. The second-largest eigenvalue  $\zeta$  and the fit of the survival times give the same decay rate, and the average survival time and (3) give the same decay rate, which confirm two conclusions. First, the second-largest eigenvalue  $\zeta$  is indeed the most dominant eigenvalue and a good approximation for the *tail* of the survival time distribution for the investigated range of  $\tau$ , in agreement with Artalejo's result [19]. Figure 1(a) shows the histograms and fitted exponential functions for four values of the effective infection rate. These results confirm that the survival time of an SIS virus in the complete graph is exponentially distributed and that the decay rate in the complete graph is indeed given by (5) and (6). For large  $\tau$ , much of the weight of the distribution is in the tail so the average of the distribution of  $T$  is well approximated by  $\zeta^{-1}$ , i.e.,  $E[T] \approx |\zeta^{-1}|$ . A more detailed analysis of the survival time distribution, including the start of the distribution, is presented in Sec. II A1. Since the average survival time starting from the quasistationary or metastable state is given by Artalejo [19] as  $E[T] = |\zeta^{-1}|$ , we expect that for small values of  $\tau$  starting from the all-infected state and from the quasistationary state will lead to increasingly different average survival times.

Second, employing the exact result of the average hitting time of the absorbing state, Appendix C proves  $E[T] = F(\tau)$  and, hence, that  $F(\tau)$  in (6) is the exact average survival time in the complete graph  $K_N$ . When the order term in (5) becomes significant in magnitude, the distribution of the survival time is no longer well described by the tail alone, which corresponds to the value of  $\tau$  in Fig. 1(a) where the four different approaches diverge.

Figure 2 shows the average survival time of the SIS process in a complete graph on 100 nodes as a function of

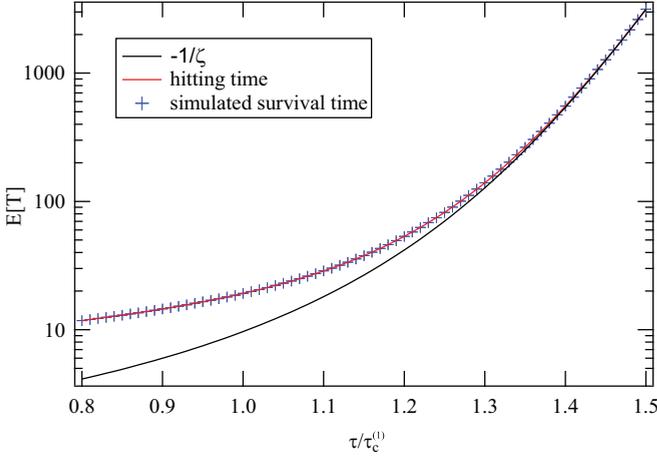


FIG. 2. (Color online) Average survival time in a complete graph of 100 nodes as a function of the normalized effective infection rate, computed via the hitting time, simulations and approximated by  $|\zeta|^{-1}$ .

the normalized effective infection rate  $\tau/\tau_c^{(1)}$  computed via a recursion for the hitting time Eq. (C6) and by simulations. The N-intertwined mean-field approximation (NIMFA) epidemic threshold  $\tau_c^{(1)}$  is defined as the reciprocal of the largest eigenvalue  $\lambda_1$  of the adjacency matrix and is a lower bound of the real epidemic threshold [23,24], i.e.,  $\tau_c^{(1)} \leq \tau_c$ . The simulations are in perfect agreement with the average hitting time. The approximation  $E[T] \geq |\zeta|^{-1}$  is also shown in Fig. 2 and is very accurate for large effective infection rates.

### 1. Survival time distribution around the epidemic threshold

Besides the expectation  $E[T]$ , the uniformized embedded Markov chain enables us to derive the complete probability density function (pdf) of the survival time  $T$ , measured in units of 1 (as mentioned above). Let  $P^k$  be the  $k$ -step transition probability matrix of the uniformized embedded Markov chain [12] describing the SIS process on the complete graph, and then  $(P^k)_{N+1,1}$  is the probability that the virus has died out after  $k$  steps, given that initially all nodes are infected. The probability that the virus has not died out yet after  $k$  steps is then given by  $\Pr[T \geq k] = 1 - (P^k)_{N+1,1}$  and  $\Pr[T = k] = \Pr[T \geq k] - \Pr[T \geq k + 1]$ . Figure 3 shows the average survival time distribution in a complete graph with 100 nodes for values of the effective infection rate  $\tau$  at different fractions of the NIMFA epidemic threshold ranging from 0.8 to 1.2. The curve for  $\tau = \tau_c^{(1)}$  is indicated by the dashed black line. Figure 3 illustrates that the curves, both above and below the epidemic threshold, are similarly shaped. For values of the effective infection rate  $\tau$  above  $\tau_c^{(1)}$ , the peak of the distribution reduces and moves slightly towards the right, while the exponential tail becomes less steep. The exponential tails of the distributions are drawn on lin-log scale in the inset of Fig. 3. The inset shows that nothing special happens around the threshold in terms of the survival time distribution of the SIS process. The slopes of the exponential tail become less and less steep for increasing infection rates, but no transition is observed around the epidemic threshold.

When the effective infection rate  $\tau$  is reduced to zero, only curing can occur. The probability density function of the

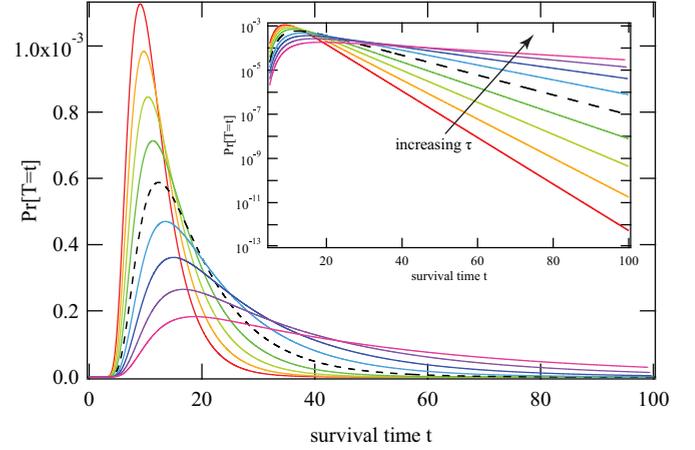


FIG. 3. (Color online) Survival time distribution of the SIS process in a complete graph of 100 nodes for different values of  $f = \frac{\tau}{\tau_c^{(1)}}$  ranging from 0.8 to 1.2. The curve for  $\tau = \tau_c^{(1)}$  is indicated by the dashed black line. The inset shows the same data on a lin-log scale to emphasize the exponential tail of all distributions both above and below the epidemic threshold.

survival time  $T$  of the SIS process for  $\tau = 0$ , denoted by  $X = T_{\tau=0}$ , is given by  $f_X(x) = \delta I(1 - e^{-\delta x})^{I-1} e^{-\delta x}$ , where  $I$  is the number of initially infected nodes, which is the distribution of the maximum [12] of  $I$  independent and identically distributed exponentially distributed random variables, all with mean  $\frac{1}{\delta}$ :

$$\begin{aligned} f_X(x) &= \frac{d}{dx} \Pr \left[ \max_{1 \leq m \leq I} X_m \leq x \right] = \frac{d}{dx} \prod_{m=1}^I \Pr[X_m \leq x] \\ &= \frac{d}{dx} (1 - e^{-\delta x})^I = \delta I (1 - e^{-\delta x})^{I-1} e^{-\delta x}, \end{aligned} \quad (7)$$

which may explain the bell shape in Fig. 3.

Even if the changes in the survival time distribution are subtle around the epidemic threshold, the influence on the average survival time displays “transition behavior.” Figure 4 shows the survival time of the SIS process in the complete graph around the epidemic threshold for various network sizes  $N$ . The threshold behavior, characterized by the bending

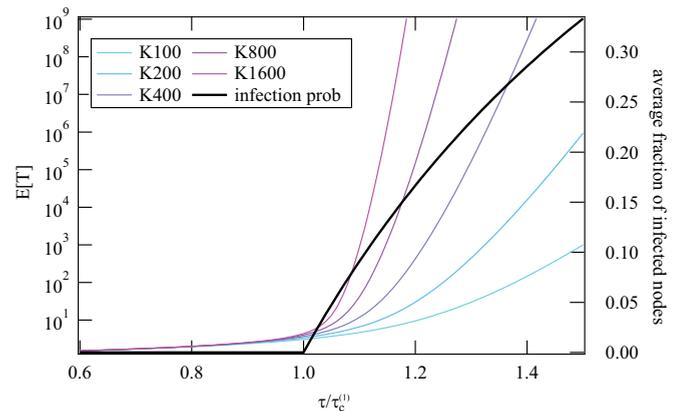


FIG. 4. (Color online) The average survival time of the SIS process in the complete graph as a function of the normalized effective infection rate  $\frac{\tau}{\tau_c}$  for various network sizes.

at  $\frac{\tau}{\tau_c^{(1)}}$  slightly above 1 and followed by a steep increase, changes with the network size  $N$ . Even for a network of 1600 nodes, the average survival time increases faster than exponential after the threshold, which agrees with the asymptotic approximation [22] of

$$F(\tau) \sim \frac{f\sqrt{2\pi} \exp\{N[\log(f) + f^{-1} - 1]\}}{\delta(f-1)^2\sqrt{N}},$$

with  $f = \tau/\tau_c$ . Figure 4 also shows the NIMFA average fraction of infected nodes in  $K_N$

$$y_\infty = 1 - \frac{1}{(N-1)\tau} = 1 - \frac{1}{\tau/\tau_c^{(1)}}$$

for  $\tau > \tau_c^{(1)} = \frac{1}{N-1}$  on the right-hand side axis for reference. With increasing network size, even small fractions  $y_\infty$  of infected nodes (for  $\tau$  just above  $\tau_c^{(1)}$ ) can last in the network for very long times. A comparison with the NIMFA solution also shows that the transition point moves closer towards  $\tau_c^{(1)}$  with increasing  $N$ , very similar to the peak in  $E[T]_{I=1}/E[T]_{I=N}$  in Fig. 7(a).

### B. The star graph

The key to reducing the  $2^N$  state space in the case of a star graph lies in realizing that the virus can only spread when the center node is infected. Moreover, all states in which the center node is either infected or healthy are identical, allowing us to reduce the number of states to  $2N$ , as illustrated in Fig. 17. The average hitting time equations (2) can also be solved for the star graph, as shown in Appendix B. Figure 5 shows the average survival time as a function of the network size for different values of the effective infection rate  $\tau \geq \tau_c^{(1)} = 1/\lambda_1 = 1/\sqrt{N-1}$ . The average survival time in Fig. 5 increases logarithmically in  $N$  for relatively small effective infection rates. The apparent stabilization of  $E[T]$  is an artifact caused by the fact that the NIMFA epidemic threshold  $\tau_c^{(1)}$  is wrong [14] by a factor  $\frac{\tau_c}{\tau_c^{(1)}} = \sqrt{\frac{1}{2} \log(N) + \frac{3}{2} \log \log(N)}$ . Hence, for sufficiently large  $N$ , the value  $m\tau^{(1)}$  for any integer

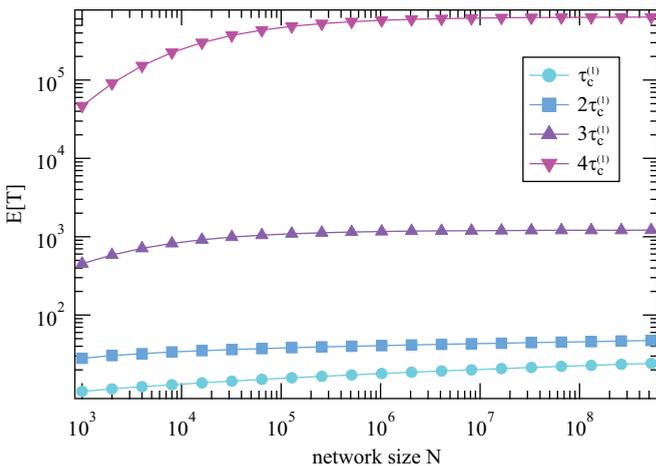


FIG. 5. (Color online) Average survival time on a star graph as a function of the size of the graph for various fractions of the NIMFA epidemic threshold  $\tau_c^{(1)}$ .

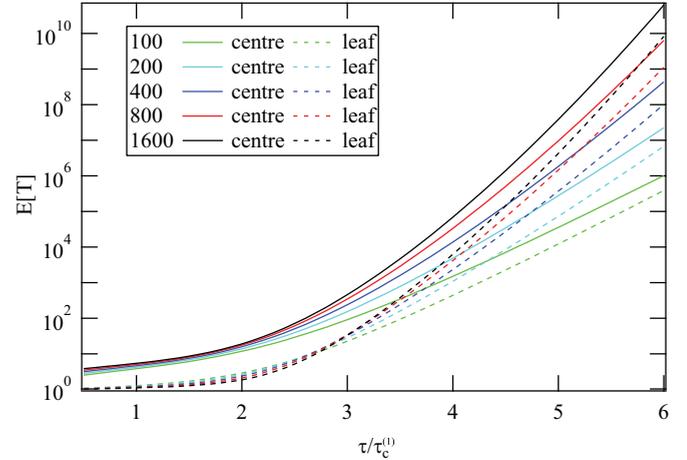


FIG. 6. (Color online) Average survival time of the SIS process starting from a single infected node in a star graph as a function of the normalized effective infection rate for various network sizes. For each size two curves are shown: a solid one indicating the process started at a center node and a dashed one indicating the process started at a leaf node.

$m$  drops below the epidemic threshold  $\tau_c$ , where  $E[T] = O(\log N)$ . In Sec. II C an approximation to the epidemic threshold in  $K_N$  and  $K_{1,N}$  is determined via the survival time, and the same correction factor for the NIMFA threshold is found.

Figure 6 shows the expected survival time starting from a single infected node in the star graph as a function of the normalized effective infection rate  $\tau/\tau_c^{(1)}$ . Obviously, there is a difference in the star graph between starting an epidemic at the center node or at one of the leaf nodes. The difference in average survival time between these two starting positions decreases slowly with increasing effective infection rate  $\tau$ . Hence, Fig. 6 illustrates that if a message or virus is desired to stay in the network for a long time, injecting the hub or center of the star is the best strategy. Although intuitive, our theoretical results allow us to quantify the difference in average survival time between choosing a leaf or hub node. Moreover, since most real-world networks can be approximately considered as a set of connected large stars, we expect that injecting the hub node will also cause the virus or message to remain longer in the network on average.

In contrast to the complete graph, the infection rate in the star graph must be a few times larger than  $\lambda_1^{-1}$  to stay in the network for a long time. In the next section, we determine, via the hitting time, precisely how much larger than  $\lambda_1^{-1}$  the epidemic threshold is. After passing the epidemic threshold, the increase in survival time is not as explosive as in the complete graph.

### C. Epidemic threshold via survival time

Although  $E[T] = F(\tau)$ , computed from (6), gives the average hitting time of the absorbing state in  $K_N$  starting from the all-infected state, the hitting time equations (2) can be solved numerically for any initial state. Using the numerical solution of (2), we can compare the survival time of a virus starting with a single infected node to the survival time, starting

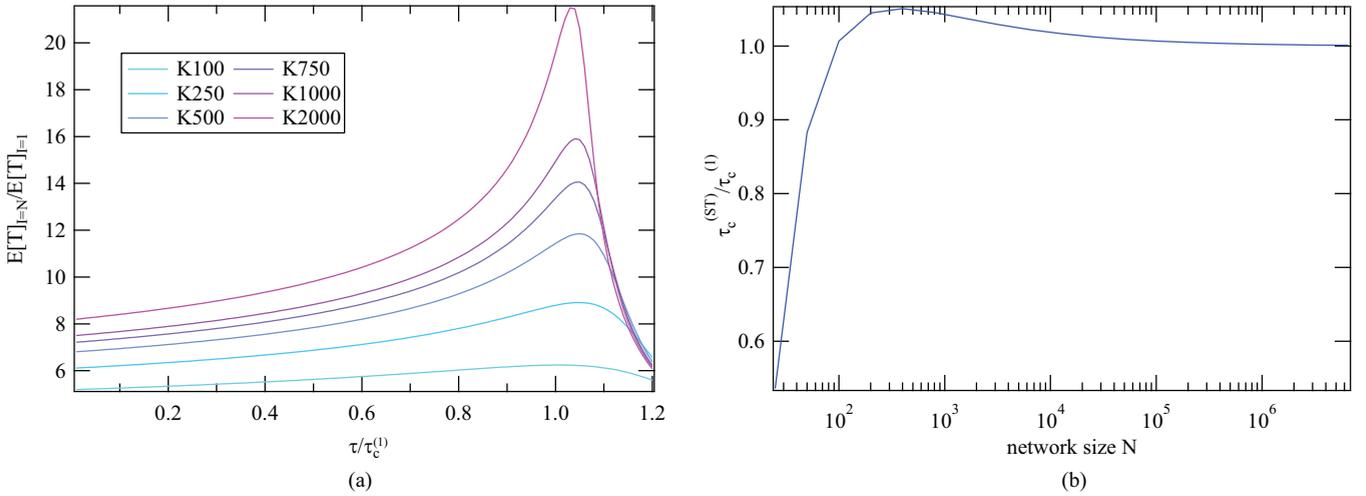


FIG. 7. (Color online) The ratio between the average survival time of the SIS process in the complete graph starting with one node infected and with all nodes infected  $\frac{E[T]_{I=N}}{E[T]_{I=1}}$  as a function of  $\tau/\tau_c^{(1)}$  for different network sizes (a). The ratio of  $\tau_c^{(ST)}/\tau_c^{(1)}$  as a function of the network size (b).

with all nodes infected. Figure 7(a) shows the ratio of the average survival time starting with a single infected node and starting in the all-infected state,  $r = E[T]_{I=N}/E[T]_{I=1}$  in  $K_N$  as a function of the normalized effective infection rate  $\tau/\tau_c^{(1)}$  for various network sizes  $N$  and is generally largest just after the NIMFA threshold  $\tau_c^{(1)} = \frac{1}{N-1}$  and slowly seems to move towards  $\tau_c^{(1)}$  with increasing  $N$ .

The peak or maximum ratio  $r$  that occurs at  $\tau = \tau_c^{(ST)}$  (ST denotes survival time) can, just as the epidemic threshold  $\tau_c$ , be interpreted as a crossover point between a regime where the curing process dominates the survival time to a regime where the infection process dominates. Left of the peak ( $\tau < \tau_c^{(ST)}$ ), the survival time is dominated by the combined curing processes of the nodes; with increasing infection rate  $\tau$ , the ratio  $r$  increases: It takes longer to cure for  $N$  nodes than just one. Right of the peak ( $\tau > \tau_c^{(ST)}$ ), however, spreading processes dominate the survival time. Above the epidemic threshold ( $\tau > \tau_c$ ), the survival time  $T$  consists of two contributions: the convergence time  $T_{I \rightarrow M}$  from the initial state with  $I$  infected nodes to the metastable state  $M$  and the remaining time  $T_{M \rightarrow A}$  from reaching the metastable state  $M$  towards absorption  $A$ , where the virus is extinct. Hence, the ratio  $r = E[T]_{I=N}/E[T]_{I=1}$  can be written as

$$r = \frac{E[T_{I=N \rightarrow M}] + E[T_{M \rightarrow A}]}{E[T_{I=1 \rightarrow M}] + E[T_{M \rightarrow A}]} = \frac{1 + \frac{E[T_{I=N \rightarrow M}]}{E[T_{M \rightarrow A}]}}{1 + \frac{E[T_{I=1 \rightarrow M}]}{E[T_{M \rightarrow A}]}}.$$

The larger  $\tau > \tau_c$ , the closer the initial state  $I = N$  lies to the number of infected nodes in the metastable state  $M$  and the smaller the time to reach  $M$  from  $I = N$ . The converse holds for the initial state  $I = 1$ . Hence,  $r$  rapidly becomes smaller. For  $\tau > \tau_c$ , the spreading times  $E[T_{I=1 \rightarrow M}]$  and  $E[T_{I=N \rightarrow M}]$  are far smaller than the average duration of the metastable state  $E[T_{M \rightarrow A}]$ . With increasing effective infection rate  $\tau$ , the spreading time decreases, while  $E[T_{M \rightarrow A}]$  increases more than exponentially, forcing  $r$  quickly to 1.

We show that the peak in  $r = E[T]_{I=N}/E[T]_{I=1}$  at  $\tau = \tau_c^{(ST)}$  provides an accurate approximation of the epidemic

threshold, thus  $\tau_c \approx \tau_c^{(ST)}$  for large  $N$ . Time-based properties of the SIS process, instead of the fraction of infected nodes, to identify the epidemic threshold were recently discussed by Boguna *et al.* [25]. The ratio  $\tau_c^{(ST)}/\tau_c^{(1)}$  as a function of the network size  $N$  is plotted in Fig. 7(b). For large  $N$  in  $K_N$ , both the NIMFA epidemic threshold  $\tau_c^{(1)}$  and the threshold  $\tau_c^{(ST)}$  approximated from the survival time are the same. For very small networks, however, the NIMFA threshold is larger than the “survival time threshold”  $\tau_c^{(ST)}$ , which may seem surprising, since the NIMFA threshold is proved to be a lower bound for the exact epidemic threshold [24]. For small network sizes  $N$ , the epidemic threshold is not precisely defined, because there is a relatively broad transition region in  $\tau$ . Only in large networks do sharper transitions appear and, ultimately, a zero-one sharp phase transition that precisely determines the epidemic threshold. Similarly, our newly defined approximation of the epidemic threshold is increasingly accurate and close to the “real” threshold for large graphs.

Similarly to the complete graph,  $\tau_c^{(ST)}$  can also be determined numerically for the star graph. In Fig. 8, the ratio  $\tau_c^{(ST)}/\tau_c^{(1)}$  is compared to the ratio  $\tau_c/\tau_c^{(1)}$  as deduced in Ref. [14]: For large  $N$ ,

$$\tau_c/\tau_c^{(1)} \sim \sqrt{1/2 \log N + 3/2 \log \log N}$$

for the star graph and

$$\tau_c/\tau_c^{(1)} \sim 1 + \frac{c}{\sqrt{N}} \quad \text{with } c > \frac{10}{3}$$

for the complete graph. Figure 8 illustrates the good agreement, suggesting that  $\tau_c \approx \tau_c^{(ST)}$  for large  $N$ .

### III. THE AVERAGE SURVIVAL TIME IN OTHER GRAPH TYPES

In this section, we use simulations and numerical methods to investigate the survival time in general graphs. The simulation results are obtained as follows. Starting from the initial state with all nodes infected, we simulate the SIS process and sample the number of infected nodes. The sample points are

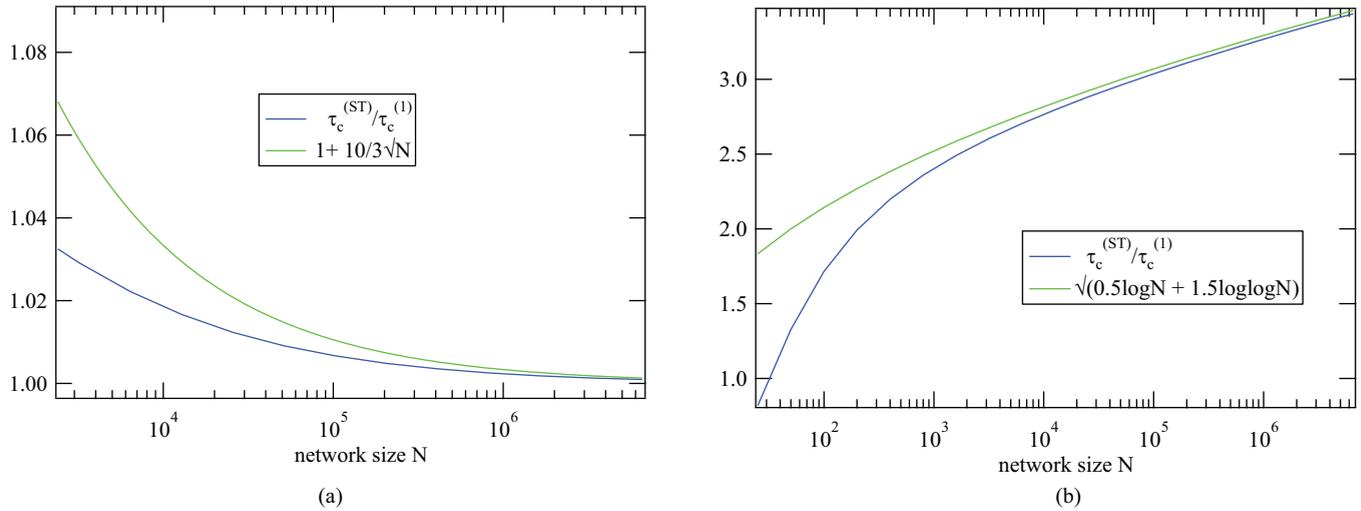


FIG. 8. (Color online) The ratio of  $\tau_c^{(ST)}/\tau_c^{(1)}$  compared to the scaling in Ref. [14] as a function of the network size for the complete graph (a) and the star graph (b).

placed at times  $(\frac{3}{2})^k$  for  $-20 \leq k \leq 32$  ranging from  $3 \times 10^{-4}$  to  $4.5 \times 10^5$ . Samples are taken regardless of the viral state of the network. If the virus dies out before the end of the simulation, then all subsequent samples for that particular run will be 0. The results shown in this section are averaged over 10 000 runs.

Figure 9 shows that the average survival time in the Erdős-Rényi random graph and the square lattice is exponentially distributed, in agreement with Markov theory. After the initial quick drop in the number of infected nodes, the decay of the number of infected nodes are straight lines in the lin-log plot, indicating the exponential decay. The insets in Fig. 9 show the expected number of infected nodes as a function of time on a logarithmic time axis. The compression of large time values makes the plateaulike metastable state visible. Both the Erdős-Rényi graph and the square lattice show similar behavior. Indeed, all graph types are expected to show similar

metastable state behavior; the only two differences between different graph types are the number of infected nodes in the metastable state as a function of the effective infection rate  $\tau$  and the duration of the metastable state: the survival time.

We first show in Fig. 10(a) the difference in the expected survival time for a fixed fraction of infected nodes in the metastable state  $y_\infty$  for various different graphs: the ring graph, the path graph, the square lattice (grid), the complete graph, the connected Erdős-Rényi random graph with link probability  $2 \log(N)/N$ , a grown preferential attachment graph with  $m = 4$  new links per node, and a star graph. All graphs have 64 nodes and the effective infection rate  $\tau$  is chosen in such a way that the average fraction of infected nodes in the metastable state is 0.45 and the same for all graphs. Figure 10(a) shows that the ring and path graphs are the two graph types that can sustain an outbreak the shortest, whereas the star graph can sustain an outbreak the longest. A virus

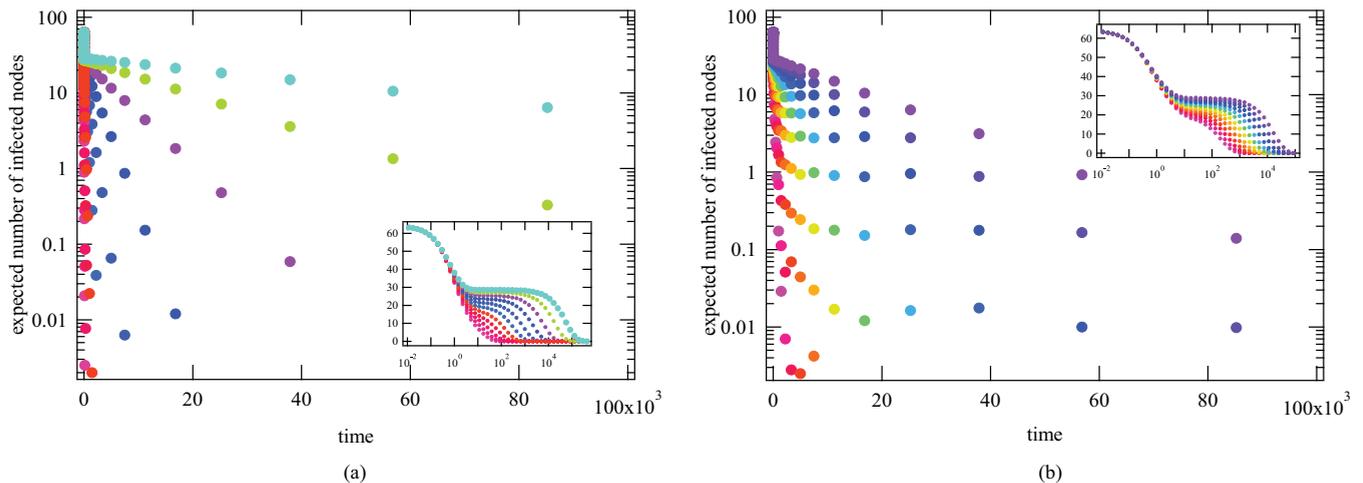


FIG. 9. (Color online) The expected number of infected nodes as a function of time for two graphs of 64 nodes for various values of the infection rate  $\beta$ . The curing rate  $\delta = 1$  for all simulations. The insets in the plots show the expected number of infected nodes with a logarithmic time axis highlighting the plateaulike metastable state of the SIS process. (a) An Erdős-Rényi graph of 64 nodes for  $\beta$  ranging from 0.13 to 0.23. (b) A square lattice of 64 nodes for  $\beta$  ranging from 0.55 to 0.67.

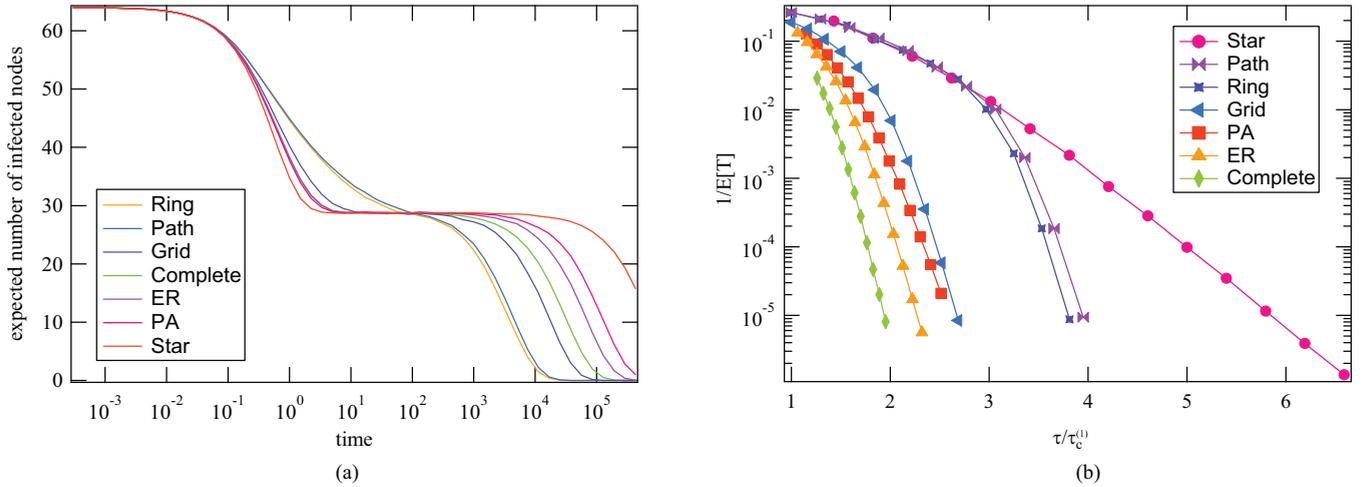


FIG. 10. (Color online) Expected survival time in general graphs. (a) The expected number of infected nodes in various graph types for an effective infection rate  $\tau$  that is chosen to have 45% nodes infected in the metastable state. (b) The exponents of the tail of the survival time distribution for various graphs of size  $N = 64$  as a function of  $\tau/\tau_c^{(1)}$  where  $\tau_c^{(1)} = \lambda_1^{-1}$  is the NIMFA epidemic threshold.

survives longest in the star graph because, to achieve 45% infected nodes in the metastable state, the effective infection rate  $\tau$  is the smallest for the star graph among all above compared graphs. All simulations start with all nodes infected, so the number of infected nodes first drops to the level of the metastable state. Figure 10(a) illustrates that the longer the process stays in the metastable state, the quicker it reaches the metastable state from the all-infected state. However, the differences are not nearly as large as the difference in the survival time for most graphs, with the exception of the ring and path graph.

#### A. Extreme graphs: Complete and star graphs

Figure 10(b) shows the decay rate  $1/E[T]$  for seven different graph types as a function of the normalized effective infection rate  $\tau/\tau_c^{(1)} > 1$ . Clearly, the complete graph has the steepest decline in decay rate. In the star graph, the SIS process needs a much higher normalized effective infection rate to survive as long as in the complete graph. The path and ring graphs follow the star graph for smaller values of the effective infection rate but then branch off for larger values of the infection rate. The point where the path and ring graph branch off from the star graph is likely the location of the epidemic threshold in these two graphs. In Sec. III B, we will show that the epidemic threshold for the ring graph of this size lies around 2.5–2.8. Above the epidemic threshold, the ring and path graphs can sustain an outbreak longer than the star graph. Figure 10(b) also suggests that, as the star and the complete graph are two extreme graphs, the average survival time  $E[T]$  as a function of the normalized effective infection rate  $\tau/\tau_c^{(1)}$  in all other graphs is expected to lie between these two graphs. In the extreme cases of infection rates either zero or infinity, every graph type shows the same behavior, but in between these two extremes the observed behavior in Fig. 10(b) holds. Combining the information from Figs. 10(a) and 10(b) leads to the conclusion that the star graph is the graph where a virus stays active the *longest* when the metastable state fraction of infected nodes is kept at 45% but the *shortest* when the

normalized effective infection rate  $\tau/\tau_c^{(1)}$  is kept constant. In part, the latter is caused by the fact that the NIMFA threshold for the star graph is not very accurate.

#### B. The ring graph

Mountford *et al.* [26] have shown that the survival time of a contact process on a tree with bounded degree is exponential in the number of nodes, which implies that the survival time on a path graph is exponential in  $N$ . As the ring graph is a path graph with one extra link, the survival time for a given effective spreading rate  $\tau$  is expected to be *at least* equal to that of the path graph.

The method, explained in Sec. II C, to approximate the epidemic threshold from the survival time, is applied in Fig. 11 to the ring graph of  $N = 500$  and  $N = 1000$  nodes. For  $N = 500$ , the “survival time” epidemic threshold  $\tau_c^{(ST)}$  lies

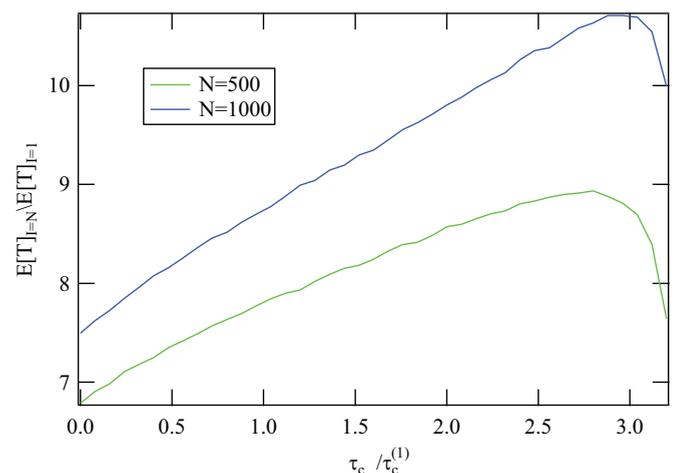


FIG. 11. (Color online) The ratio between the average survival time of the SIS process in the ring graph starting with one node infected and with all nodes infected  $\frac{E[T]_{I=N}}{E[T]_{I=1}}$  as a function of  $\tau/\tau_c^{(1)}$  for different network sizes.

at approximately  $2.8\tau_c^{(1)}$ , while, for  $N = 1000$ ,  $\tau_c^{(ST)} \approx 3\tau_c^{(1)}$ , although the NIMFA epidemic threshold  $\tau_c^{(1)} = \frac{1}{\lambda_1} = 0.5$  is constant for any ring. We infer from Fig. 11 that, for a ring,  $\tau_c = \frac{f_{ring}(N)}{\lambda_1}$ , where  $f_{ring}(N)$  is a slowly increasing function of  $N$ . Unfortunately, the determination of  $f_{ring}(N)$  via  $\tau_c^{(ST)}$  for larger graphs is computationally very expensive.

**IV. NON-MARKOVIAN SPREADING**

The exact results for the complete graph  $K_N$  and the star graph  $K_{1,N}$  that underpin the results in Sec. II are only attainable because the infection and spreading processes are independent Poisson processes. The analysis of non-Markovian epidemic spreading is complicated by the absence of the memoryless property of the exponential distribution. Yet it has received attention in the literature [27–30]. In this section, we show the influence on the survival time of a non-Poissonian infection process. The NIMFA infection probability in the metastable state is derived in Ref. [31],

$$E[M](1 - v_{i\infty}) \sum_{j=1}^N a_{ij} v_{j\infty} = v_{i\infty},$$

where  $E[M]$  is the expected number of infection attempts during an infectious period of a node,  $a_{ij}$  are the components of the adjacency matrix  $A$  of the graph, and  $v_{i\infty}$  is the probability that node  $i$  is infected in the metastable state. In the case of exponentially distributed interarrival times between infection and spreading events,  $E[M]$  is given by  $\tau = \frac{\beta}{\delta}$ , while the general expression for  $E[M]$  is deduced in Ref. [31]. The NIMFA steady-state infection probability  $v_{i\infty}$  is, however, no longer an upper bound, as shown in Ref. [24] and further illustrated in Ref. [32]. The average survival time  $E[T]$  for non-Markovian SIS processes is deduced by simulations.

In order to investigate the effect of heavy-tailed distributions for the infection times  $T_I$  on the survival time, we replace the exponential distribution with a Weibull distribution given by

$$f_{T_I}(x) = \frac{\alpha}{b} \left(\frac{x}{b}\right)^{\alpha-1} e^{-(x/b)^\alpha}, \tag{8}$$

where  $\alpha$  is generally called the shape parameter and  $b$  the scale parameter. For  $\alpha = 1$ , the Weibull distribution reduces to the exponential distribution. For  $\alpha < 1$ , the distribution is heavy-tailed and reduces to a Zipf distribution for  $\alpha = 0$ . For large values of  $\alpha$ , the tail falls off exponentially and for  $\alpha \rightarrow \infty$  the Weibull distribution reduces to a Dirac function. The shape of the Weibull distribution as a function of  $\alpha$  is shown in the inset of Fig. 13(a).

The difference between the Markovian and non-Markovian SIS process is most visible in the distribution of the infection attempts over an infectious period of a node. Even when the average number of infection attempts during an infectious period  $E[M]$  is kept constant, the steady-state fraction of infected nodes and the survival time of the process change because of the timing of infection attempts relative to the curing time of a node. In this section, we only change the distribution of the infection events and keep the curing time of a node exponentially distributed.

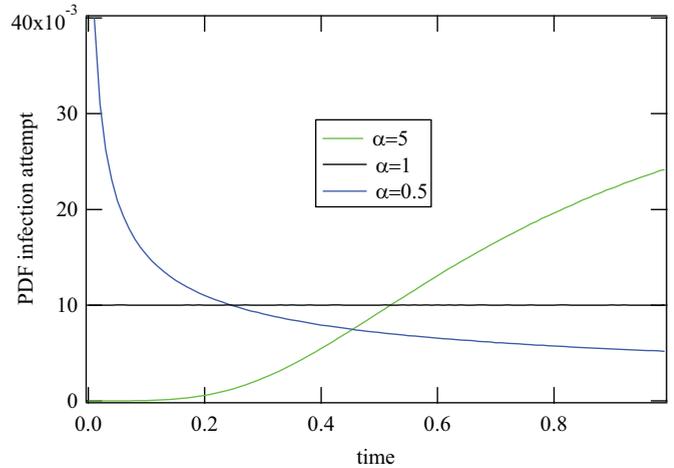


FIG. 12. (Color online) Distribution of the infection attempts over a link normalized to the infectious period of the source node.

Figure 12 shows the simulated time distribution of infection attempts over a single link, normalized to the infectious period of the source node. Figure 12 is obtained as follows: We store the occurrence time of each infection attempt, relative to the beginning of the infectious period of the node, and normalize that occurrence time to the duration of the infectious period of the node. For exponential interarrival times, given that one event has happened during an interval, its occurrence time is uniformly distributed over that time interval [12]. For other distributions, however, the occurrence of events is *not* uniformly distributed over an interval. Indeed, in Fig. 12, the exponential case ( $\alpha = 1$ ) is uniform, as expected, but for  $\alpha < 1$  events tend to be early and for  $\alpha > 1$  late. We believe that the deviation from uniform infection attempts over an infectious period makes the non-Markovian SIS different in its behavior from the classical, Markovian SIS epidemic.

The SIS process only stays long in a network if nodes become reinfected after they cure. Therefore, the timing of the infection attempts greatly influences the process. When the virus spreads from a source node early in its infectious period, it is less likely that the newly infected neighbor will re infect the source node after it cures, because the newly infected neighbor is also likely to spread early and the infectious state of the source node has most likely not changed yet. On the other hand, if the source node spreads the infection towards the end of its infectious period, it is likely to be cured and susceptible again if the newly infected neighbor spreads the infection back.

Figure 13(a) shows the survival time distribution for  $\alpha$  ranging from 0.5 to 1.5 and a constant value for the expected number of infection attempts during an infectious period  $E[M] = 0.014$  starting from the all-infected state.

For small values of the shape parameter  $\alpha$ , the survival time distribution falls off much quicker than for the exponential case (shown in black), due to the early infection attempts, explained in Fig. 12. The virus does not succeed in re infecting nodes that are cured and, thus, dies out quicker. For larger values of  $\alpha$ , the opposite happens. The survival time distribution falls off slower than in the exponential case.

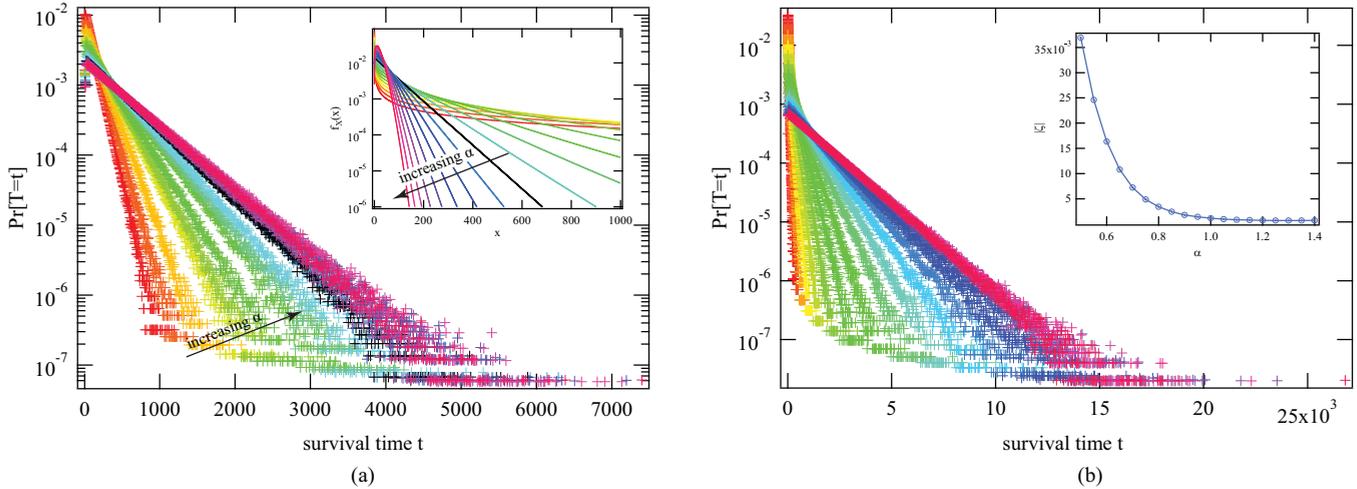


FIG. 13. (Color online) Survival time as a function of the shape parameter  $\alpha$  in the complete graph (a) and the star graph (b). (a) Distribution of the survival time for Weibull distributed interarrival times for various values of the shape parameter  $\alpha = \{0.5, 1.3\}$  and  $E[M] = 0.014$  in the complete graph of 100 nodes. The exponential case ( $\alpha = 1$ ) is indicated in black. Inset: The Weibull distribution. (b) Survival time distribution in a star graph of 100 nodes for different values of the shape parameter  $\alpha = \{0.5, 1.4\}$ . Scale parameter  $b$  is chosen such that on average 25% of the nodes are infected in the metastable state. Inset: The rates of the exponential tails as a function of  $\alpha$ .

The same influence of the shape parameter  $\alpha$  on the average survival time, observed in the complete graph, is also visible in the star graph, as shown in Fig. 13(b), where the survival time distribution for various values of  $\alpha$  is drawn in a star graph of 100 nodes. The inset shows the decay rate, the exponent of the tails, as a function of the shape parameter  $\alpha$ . In this case, we have kept the steady-state fraction of infected nodes constant at 25% by changing the scale parameter  $b$  in (8) appropriately.

The effect of very small values for the shape parameter  $\alpha$  is shown in Fig. 14(a), where the survival time distribution is shown for a complete graph of 100 nodes for various values of  $\alpha$ . The black curve indicates the survival time of the virus, when the infection rate is zero, in which case nodes can only cure. Since all nodes are infected in the initial state, the

black curve shows the the probability distribution (7) of the maximum of 100 exponential independent and identically distributed random variables with rate  $\delta$ . Figure 14(a) illustrates that, for small values of the shape parameter  $\alpha$ , the distribution of the survival time resembles that of the no spreading case: All spreading attempts occur shortly after infection, as shown in Fig. 12, after which all nodes cure again as if there is no infection taking place.

The effect of relatively large values for the shape parameter  $\alpha$ , on the other hand, is shown in Fig. 14(b) for three different graph types: the complete graph, the star graph, and the square lattice. For small values of  $\alpha$ , the expected survival time of the virus is short, as mentioned above. As the shape parameter  $\alpha$  increases, so does the expected survival time. However, the

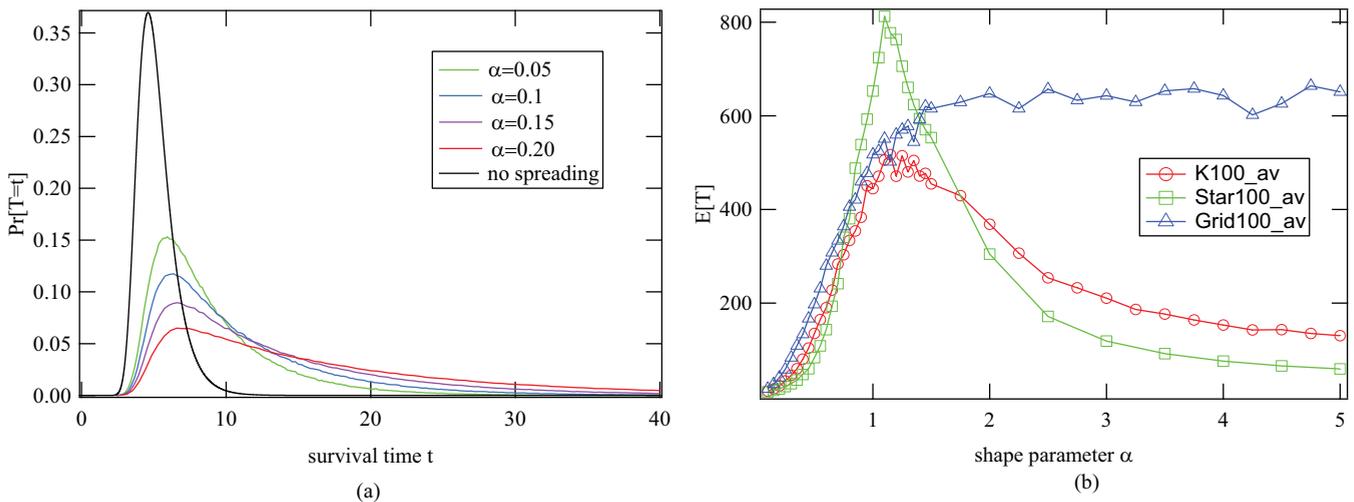


FIG. 14. (Color online) Survival time as a function of shape parameter  $\alpha$ . (a) Survival time distribution in the complete graph of 100 nodes for small values of the shape parameter  $\alpha$ . The scale parameter  $b$  is chosen such that  $E[M]$  is constant. (b) Average survival time as a function of the shape parameter  $\alpha$  in the complete graph, the star graph, and the square lattice, all with  $N = 100$  nodes. The scale parameter  $b$  is chosen such that on average 25% of the nodes are infected in the metastable state.

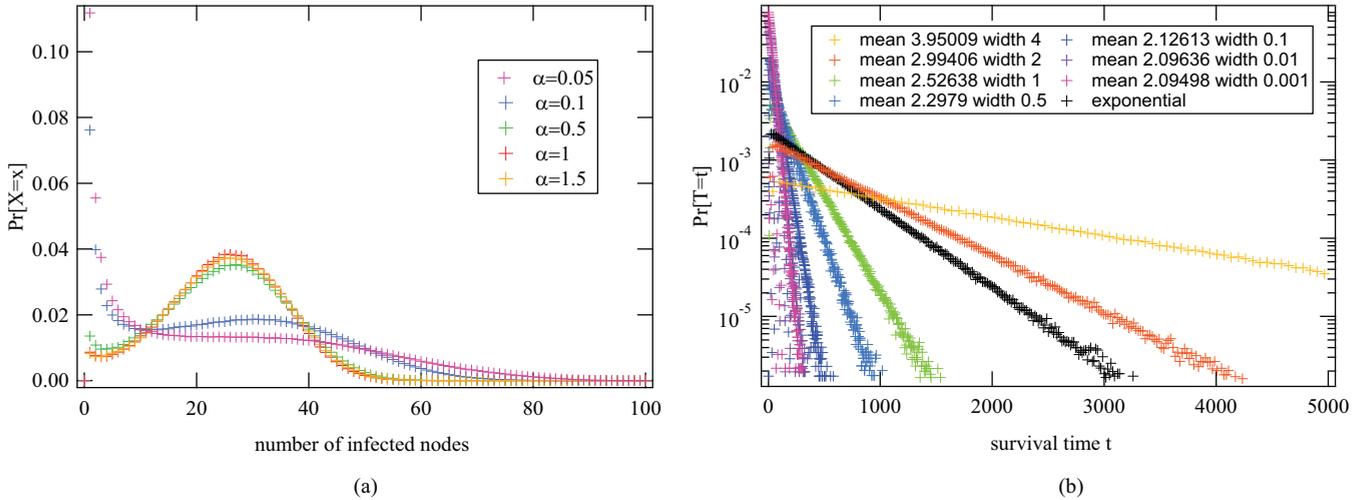


FIG. 15. (Color online) (a) The average number of infected nodes in a square grid of 100 nodes for various values of the shape parameter  $\alpha$  and a scale parameter  $b$  chosen in such a way that the average is constant. The virus is prevented from dying out. (b) Survival time distribution of the SIS process with uniformly distributed cure and infection times for different values of the width of the infection rate. The mean of the infection rate is scaled to have on average 25% of the nodes infected in the metastable state. The exponential case is shown in black for reference.

expected survival time peaks around  $\alpha = 1.1$  for the star graph and complete graph and then starts to decrease. This might be caused by a form of synchronization in the infection times. The larger the shape parameter  $\alpha$  becomes, the narrower the infection time distribution becomes. As a result, infections occur in an increasingly small interval after the source node becomes infected. When the infectious periods of the nodes are aligned in time, the chance that the virus dies out increases. Dying out is more likely in the star and complete graphs because of the very small average hop count. In the case of the square grid, the shortening of the survival time with increasing  $\alpha$  is not observed; the expected survival time does not peak but stabilizes. Figure 14(b) illustrates that the peak is much sharper in the star graph than in the complete graph.

The influence of a small shape parameter  $\alpha$  as described above on the distribution of the number of infected nodes in the metastable state is shown in Fig. 15(a), which shows the distribution of the number of infected nodes for various values of the shape parameter  $\alpha$ . To simulate the metastable distribution of infected nodes, we simulate a modified SIS (MSIS) process [14] where the absorbing state is removed and take the steady state of the MSIS process to be the metastable state of the original SIS process. For smaller values of the shape parameter  $\alpha$ , the distribution peaks around a single infected node and is relatively flat after that peak. States with a large number of infected nodes are not unlikely, indicating that the virus quickly spreads as a result of the nodes spreading early during the infectious period and then dies out again.

To show the effect of both late infection times and synchronization, we use uniformly distributed cure and infection times. The width and mean of the cure times is kept constant, but the width of the infection time distribution is varied while the mean is scaled to keep the average fraction of infected nodes  $y_\infty$  constant. Because both the curing and infection times are uniformly distributed, the interval in which infection attempts can fall lies between the start of the infection distribution and

the end of the curing distribution. The values used in Fig. 15(b) imply an interval for infection attempts between  $[1.95, 2.1]$  and  $[2.09, 2.1]$ . Figure 15(b) shows the survival time distribution in the complete graph of 100 nodes for various values of the width of uniform infection time distributions. The average survival time distribution also has an exponential tail when *both* the infection and the curing times are nonexponential. Figure 15(b) shows that the late infection times lead to longer survival times than for exponential infection times (shown in black) but that for the narrower intervals the expected survival time is shorter. For the narrowest width of the infection time distribution, any infection attempt lies in an interval only 0.01 time units wide, which leads to synchronized behavior. Although late infections lead to longer expected survival times, synchronized spreading leads to shorter infection times.

## V. CONCLUSION

In this paper, the survival time  $T$  of an SIS process on a network has been investigated. We have derived exact equations for the average survival time in the complete graph  $K_N$  and the star graph  $K_{1,N}$ , using the hitting time of a uniformized embedded Markov chain. For the SIS process on the complete graph  $K_N$ , we have proved that the average survival time  $E[T] = F(\tau)$ , which is the first term in the Lagrange series (5) of the second largest eigenvalue  $\zeta$  of the infinitesimal generator  $Q$  of the continuous-time Markov SIS process. Numerically efficient expressions are deduced such as the recursive (3).

Relying on the hitting time formalism in Markov theory, a new method to accurately approximate the epidemic threshold from the survival time is presented: The epidemic threshold occurs at the effective infection rate  $\tau_c^{(ST)}$  that corresponds to the peak in  $\frac{E[T]_{l=N}}{E[T]_{l=1}}$ .

Simulations have shown that graphs other than  $K_N$  and  $K_{1,N}$  also exhibit exponentially distributed survival times. When the steady-state fraction  $y_\infty$  of infected nodes is kept

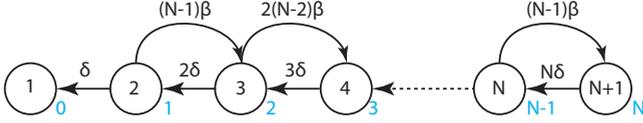


FIG. 16. (Color online) State diagram and transitions of the SIS process in the complete graph. The state numbers are indicated inside the states, whereas the number of infected nodes in each state is indicated in blue next to the state. The transition rates to states with more infected nodes (towards the right) depend on the number of healthy and infected nodes and the infection rate  $\beta$ , while the transition rates to states with fewer infected nodes (towards the left) depend on the number of infected nodes and the curing rate  $\delta$ .

constant (e.g., at 45%), the star graph sustains the infection the longest, whereas the ring graph sustains the infection the shortest. As a function of the normalized effective infection rate  $\tau/\tau_c^{(1)}$ , however, the virus dies out the quickest in the star graph. These results are caused by the accuracy of the NIMFA epidemic threshold  $\tau_c^{(1)} = \frac{1}{\lambda_1}$ , used in the normalization  $\tau/\tau_c^{(1)}$  and underline the difficulty of finding an unbiased way to compare the SIS process in different graph types.

Finally, in non-Markovian spreading, the virus dies out quicker for heavy-tailed infection time distributions, compared to exponential infection times, for the same number of expected infection attempts during an infectious period. For heavy-tailed distributions, these infection attempts tend to occur early in the infectious period, thereby minimizing the probability that a source node can be reinfected by a neighbor node, because the source node is most likely still infected,

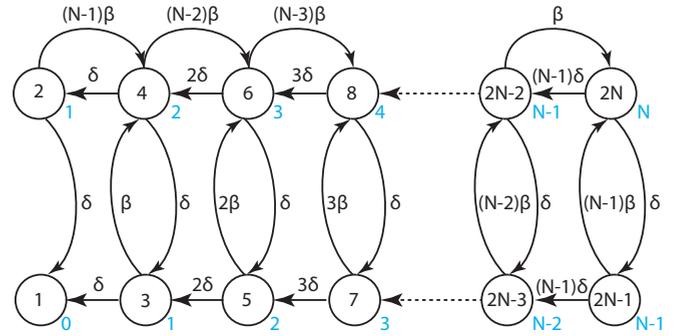


FIG. 17. (Color online) Markov Chain representation of the SIS virus process in the star graph. In the odd state the center node is not infected, whereas in the even states the center node is infected. The number of infected nodes is indicated in blue next to the states.

when a neighbor tries to reinfected the source node. As a result, the infection initially spreads quickly through the network and then dies out again. Many real-world dynamic processes are non-Markovian [30], which motivates us to understand the differences between non-Markovian and Markovian SIS.

#### APPENDIX A: SURVIVAL TIME IN $K_N$

The SIS process on a complete graph can be described [12] by a continuous-time Markov chain, as illustrated in Fig. 16. States are numbered from 1 to  $N$ , while the number of infected nodes in state  $i$  is  $i - 1$ . The infinitesimal generator is given by:

$$-Q = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 + (N-1)\tau & -(N-1)\tau & 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & 2 + 2(N-2)\tau & -2(N-2)\tau & 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & 3 + 3(N-3)\tau & -3(N-3)\tau & 0 & 0 & 0 \\ 0 & 0 & 0 & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -N & N & 0 \end{bmatrix}.$$

System (2) can be solved using Gaussian elimination, that is, by reducing the augmented matrix  $[-Q|b]$  to row echelon form. Because we are interested in the worst-case scenario, we only need to know the value of  $w_{N+1}$ , which corresponds to the hitting time of the absorbing state starting from the all-infected state. In Gaussian elimination, the system matrix is reduced to an upper-triangular matrix. Reducing all nonzero subdiagonal elements from column  $k$  for  $k \geq 2$  to zero is achieved by adding  $\frac{k}{d_k}$  times row  $k$  to row  $k + 1$ :

$$r_{k+1} \rightarrow r_{k+1} + \frac{k}{d_k} r_k,$$

where  $r_k$  is row  $k$  in  $[-Q|b]$ . The multiplication factor  $d_k$  can be found recursively for  $k > 2$  as

$$d_k = (k-1)(1 + (N-k+1)\tau) - \frac{(k-1)(k-2)(N-k+2)\tau}{d_{k-1}}, \quad (\text{A1})$$

with initial conditions  $d_1 = 1$  and  $d_2 = 1 + (N-1)\tau$ . Since all row operations are performed on the augmented system matrix, the entries in  $b$  also change. The average survival time starting from the all-infected state is found by dividing the last element in  $b$  by  $w_{N+1} = d_{N+1}$ :

$$E[T] = b_{N+1}/d_{N+1}, \quad (\text{A2})$$

where  $b_k$  is given by the recursive relation for  $k > 1$  as

$$b_k = 1 + \frac{(k-1)}{d_{k-1}} b_{k-1} \quad (\text{A3})$$

and initial condition  $b_1 = 0$ . Using (A2) and (A3), Appendix C proves that the average survival time in  $K_N$  equals  $E[T] = F(\tau)$ , which we consider as a major result.

### APPENDIX B: SURVIVAL TIME IN $K_{1,N-1}$

The continuous-time Markov chain for the star graph contains  $2N$  states, as shown in Fig. 17. The Markov chain in Fig. 17 is a relabeling of the chain in Ref. [14] so the infinitesimal generator has a five-band structure. In the odd states, the center node is healthy, whereas in the even states it is infected. The number of infected nodes in each state is indicated in blue next to the state. Just as in the case of the complete graph, we solve the system  $\phi K w = b$ , where  $\phi K = -Q$  and  $b = \mathbf{u} - e_1$  is the all one vector minus the first standard basis, by reducing the augmented matrix  $[-Q|b]$  to row echelon form. The infinitesimal generator of the SIS process on the star is given by

$$-Q = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & q_2 & 0 & -(N-1)\tau & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & q_3 & -\tau & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & q_4 & 0 & -(N-2)\tau & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & q_5 & -2\tau & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 & -1 & q_6 & 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 & q_7 & \ddots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \ddots & -1 & q_8 & 0 & -\tau & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -(N-1) & 0 & q_9 & -(N-1)\tau & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -(N-1) & -1 & q_{10} & 0 \end{bmatrix}.$$

Contrary to the complete graph, odd columns have two nonzero elements below the diagonal and reducing them to zero will also lead to two nonzero elements below the diagonal of the next (even) column.

The nonzero subdiagonal entries of odd columns for  $k > 1$  are reduced to zero by two row operations:

$$r_{k+1} \leftarrow r_{k+1} + \frac{\delta}{d_k} r_k$$

$$r_{k+2} \leftarrow r_{k+2} - \frac{(k+1)\delta}{2d_k} r_k.$$

The nonzero subdiagonal entries of even columns for  $k > 2$  are also reduced to zero by two row operations:

$$r_{k+1} \leftarrow r_{k+1} - \frac{m_k k \delta}{2d_{k-1} d_k} r_k$$

$$r_{k+2} \leftarrow r_{k+2} + \frac{k\delta}{2d_k} r_k,$$

where  $d_k$  is given by:

$$\frac{k\delta}{2} + \left(N - \frac{k}{2}\right)\beta - \frac{(k/2 - 1)(N - k/2 + 1)\beta\delta}{d_{k-2}} + \frac{\delta m_k}{d_{k-1}}, \quad \text{for even } k > 2,$$

$$\delta + (N - 1)\beta, \quad k = 2,$$

$$(k - 1)(\delta + \beta)/2, \quad \text{for odd } k > 0,$$

and  $m_k$  is given, for  $k > 4$ , by

$$m_k = -\left(\frac{k}{2} - 1\right)\beta + \frac{(k/2 - 1)(N - k/2 + 1)\beta\delta m_{k-2}}{d_{k-2} d_{k-3}},$$

with initial condition  $m_4 = -\beta$ . The worst-case expected survival time in the star graph is obtained by dividing  $w_{2N} = d_{2N}$  by the last element in  $b$ :

$$E[T] = b_{2N}/d_{2N},$$

where  $b_k$  is given by

$$1 + \left(\frac{k}{2} - 1\right) \frac{\delta}{d_{k-2}} b_{k-2} + \frac{\delta}{d_{k-1}} b_{k-1} \quad \text{for } k > 1 \text{ even,}$$

$$1 + \frac{(k-1)\delta}{2d_{k-2}} b_{k-2} - \frac{(k-1)m_{k-1}\delta}{2d_{k-2}d_{k-1}} b_{k-1} \quad \text{for } k > 1 \text{ odd,}$$

and initial conditions  $b_1 = 0$ ,  $b_2 = 1$ .

Although the expression for the average survival time in the star graph is more involved than in the case of the complete graph, it enables the numerical evaluation of the average survival time in a linear time and constant space complexity.

### APPENDIX C: PROOF THAT $E[T] = F(\tau)$ FOR SIS EPIDEMICS ON $K_N$

First, we rewrite (6) as

$$F(\tau) = \sum_{j=1}^N \frac{x_j}{j}, \quad (\text{C1})$$

where  $x_j = \sum_{r=0}^{j-1} \frac{(N-j+r)!}{(N-j)!}$  and initial conditions  $x_1 = 1$ ,  $x_j = 0$  for  $j < 0$ . Now  $x_j$  obeys the recursive relation (4). Indeed,

$$\begin{aligned} x_{j+1} &= \sum_{r=0}^j \frac{(N-j-1+r)!}{(N-j-1)!} \tau^r = \tau(N-j) \sum_{r=-1}^{j-1} \frac{(N-j+r)!}{(N-j)!} \tau^r \\ &= x_j(N-j)\tau + \tau(N-j) \frac{(N-j-1)!}{(N-j)!} \tau^{-1} = x_j(N-j)\tau + 1. \end{aligned}$$

The summation (C1) is more efficient to compute numerically as it contains only  $N$  terms instead of  $\frac{N^2}{2}$  in the double sum in (6) and does not contain factorials.

We will now show that (A2) is equal to  $F(\tau)$  in (6) or in (3). As a consequence, the first term in the Lagrange series [22] of the second-largest eigenvalue  $\zeta$  of the infinitesimal generator  $Q$  is precisely the average worst-case survival time of the SIS process on  $K_N$ .

We first rewrite  $b_{N+1}$  using the recursion (A3):

$$b_{N+1} = 1 + \frac{N}{d_N} b_N = 1 + \frac{N}{d_N} \left[ 1 + \frac{(N-1)}{d_{N-1}} b_{N-1} \right] = 1 + \frac{N}{d_N} + \frac{N(N-1)}{d_N d_{N-1}} b_{N-1}.$$

Repeating this process for all  $b_i$  until  $i = 1$  results in

$$b_{N+1} = \sum_{j=0}^{N-1} \frac{N! \prod_{i=2}^{N-j} d_i}{(N-j)! \prod_{i=2}^N d_i}.$$

Finally, dividing by  $d_{N+1}$  yields, by use of (A2), the average survival time,

$$E[T] = \sum_{j=0}^{N-1} \frac{N! \prod_{i=2}^{N-j} d_i}{(N-j)! \prod_{i=2}^{N+1} d_i}. \quad (\text{C2})$$

Next we rewrite (C2) in terms of  $p_j = \prod_{i=2}^j d_i$ , which also equals

$$p_j = d_j p_{j-1} = d_j \prod_{i=2}^{j-1} d_i \quad (\text{C3})$$

and

$$d_j = \frac{p_j}{p_{j-1}}.$$

Substituting into (A1) yields

$$d_j = (j-1)(1 + (N-j+1)\tau) - \frac{(j-1)(j-2)(N-j+2)\tau p_{j-2}}{p_{j-1}}. \quad (\text{C4})$$

Substituting (C4) into (C3) leads to the following recursive expression for  $p_j$ :

$$p_j = (j-1)[1 + (N-j+1)\tau] p_{j-1} - (j-1)(j-2)(N-j+2)\tau p_{j-2},$$

which can be further simplified by first splitting and then iterating the first term:

$$\begin{aligned} p_j &= (j-1)p_{j-1} - (j-1)(j-2)(N-j+2)\tau p_{j-2} + (j-1)(N-j+1)\tau p_{j-1} \\ &= (j-1)(j-2)[1 + (N-j+2)\tau] p_{j-2} - (j-1)(j-2)(j-3)(N-j+3)\tau p_{j-3} \\ &\quad - (j-1)(j-2)(N-j+2)\tau p_{j-2} + (j-1)(N-j+1)p_{j-1} \\ &= (j-1)(j-2)p_{j-2} - (j-1)(j-2)(j-3)(N-j+3)\tau p_{j-3} + (j-1)(N-j+1)p_{j-1}. \end{aligned}$$

The first term can be expanded recursively to arrive, with  $p_2 = d_2 = 1 + (N - 1)\tau$ , at

$$\begin{aligned} p_j &= (j - 1)(j - 2)\dots 2p_2 - (j - 1)(j - 2)\dots(N - 1)\tau p_1 + (j - 1)(N - j + 1)p_{j-1} \\ &= (j - 1)![1 + (N - 1)\tau] - (j - 1)!(N - 1)\tau + (j - 1)(N - j + 1)p_{j-1} \\ &= (j - 1)! + (j - 1)(N - j + 1)\tau p_{j-1}. \end{aligned} \quad (\text{C5})$$

From (C5), we find that  $p_{N+1} = N!$  and (C2) becomes

$$E[T] = \sum_{j=0}^{N-1} \frac{p_{N-j}}{(N-j)!} = \sum_{k=1}^N \frac{p_k}{k!}. \quad (\text{C6})$$

Moreover, let  $y_k = \frac{p_k}{(k-1)!}$ , and then in (C5)

$$(j - 1)!y_j = (j - 1)! + (j - 1)(N - j + 1)\tau(j - 2)!y_{j-1}$$

and, after rearranging terms,

$$y_j = 1 + (N - j + 1)\tau y_{j-1}.$$

Comparing with (4), we observe that  $y_j$  obeys the same recursion (including the initial conditions) as  $x_j$  in (4). Due to the uniqueness of the solution of a linear difference equation,  $y_j = x_j$ , so  $E[T] = \sum_{k=1}^N \frac{x_k}{k}$ , which is precisely (6), proving that the first term in the Lagrange series of the second-largest eigenvalue  $\zeta$  of the infinitesimal generator  $Q$  is equal to the average survival time of the SIS process on  $K_N$ . As the survival time starting from the metastable state is exponentially distributed with a rate equal to  $\zeta$ , the order term in (5) is a measure for the difference between starting in the metastable state and in the all-infected state.

- 
- [1] O. Diekmann, H. Heesterbeek, and T. Britton, *Mathematical Tools for Understanding Infectious Disease Dynamics* (Princeton University Press, Princeton, NJ, 2012).
- [2] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, Epidemic processes in complex networks, *Rev. Mod. Phys.* **87**, 925 (2015).
- [3] C. Castellano and R. Pastor-Satorras, Thresholds for Epidemic Spreading in Networks, *Phys. Rev. Lett.* **105**, 218701 (2010).
- [4] R. Parshani, S. Carmi, and S. Havlin, Epidemic Threshold for the Susceptible-Infectious-Susceptible Model on Random Networks, *Phys. Rev. Lett.* **104**, 258701 (2010).
- [5] P. Van Mieghem, J. Omic, and R. Kooij, Virus spread in networks, *IEEE/ACM Transact. Network.* **17**, 1 (2009).
- [6] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos, Epidemic spreading in real networks: An eigenvalue viewpoint, in *Proceedings of the 22nd International Symposium on Reliable Distributed Systems* (2003), pp. 25–34.
- [7] P. Van Mieghem, Epidemic phase transition of the sis type in networks, *Europhys. Lett.* **97**, 48004 (2012).
- [8] S. Gómez, A. Arenas, J. Borge-Holthoefer, S. Meloni, and Y. Moreno, Discrete-time Markov chain approach to contact-based disease spreading in complex networks, *Europhys. Lett.* **89**, 38009 (2010).
- [9] E. Cator and P. Van Mieghem, Second-order mean-field Susceptible-Infecteds-Susceptible epidemic threshold, *Phys. Rev. E* **85**, 056111 (2012).
- [10] R. Pastor-Satorras and A. Vespignani, Epidemic dynamics and endemic states in complex networks, *Phys. Rev. E* **63**, 066117 (2001).
- [11] Angélica S. Mata, and Silvio C. Ferreira, Multiple transitions of the Susceptible-Infected-Susceptible epidemic model on complex networks, *Phys. Rev. E* **91**, 012816 (2015).
- [12] Piet Van Mieghem, *Performance Analysis of Complex Networks and Systems* (Cambridge University Press, Cambridge, 2014).
- [13] A. Ganesh, L. Massoulié, and D. Towsley, The effect of network topology on the spread of epidemics, *Proc. IEEE* **2**, 1455 (2005).
- [14] E. Cator and P. Van Mieghem, Susceptible-Infected-Susceptible epidemics on the complete graph and the star graph: Exact analysis, *Phys. Rev. E* **87**, 012811 (2013).
- [15] N. Blenn, C. Doerr, and P. Van Mieghem, Interactivity patterns in online social media seem log-normally distributed (unpublished, 2015).
- [16] C. Doerr, N. Blenn, and Piet Van Mieghem, Lognormal infection times of online information spread, *PLoS ONE* **8**, 1 (2013).
- [17] P. Van Mieghem and E. Cator, Epidemics in networks with nodal self-infections and the epidemic threshold, *Phys. Rev. E* **86**, 016116 (2012).
- [18] J. R. Norris, *Markov Chains*, Cambridge Series on Statistical and Probabilistic Mathematics (Cambridge University Press, Cambridge, UK, 2005).
- [19] J. R. Artalejo, On the time to extinction from quasi-stationarity: A unified approach, *Physica A* **391**, 4483 (2012).
- [20] J. A. Fill, The passage time distribution for a birth-and-death chain: Strong stationary duality gives a first stochastic proof, *J. Theor. Probab.* **22**, 543 (2009).
- [21] L. Miclo, On absorption times and Dirichlet eigenvalues, *ESAIM: Prob. Stat.* **14**, 117 (2010).
- [22] P. Van Mieghem, Decay towards the overall-healthy state in SIS epidemics on networks, [arXiv:1310.3980](https://arxiv.org/abs/1310.3980).
- [23] Piet Van Mieghem, The N-intertwined SIS epidemic network model, *Computing* **93**, 147 (2011).

- [24] E. Cator and P. Van Mieghem, Nodal infection in Markovian Susceptible-Infected-Susceptible and Susceptible-Infected-Removed epidemics on networks are non-negatively correlated, *Phys. Rev. E* **89**, 052802 (2014).
- [25] M. Boguñá, C. Castellano, and R. Pastor-Satorras, Nature of the Epidemic Threshold for the Susceptible-Infected-Susceptible Dynamics in Networks, *Phys. Rev. Lett.* **111**, 068701 (2013).
- [26] T. Mountford, J.-C. Mourrat, D. Valesin, and Q. Yao, Exponential extinction time of the contact process on finite graphs, [arXiv:1203.2972](https://arxiv.org/abs/1203.2972).
- [27] H. Hinrichsen, Non-equilibrium phase transitions with long-range interactions, *J. Stat. Mech.: Theor. Exp.* (2007) P07006.
- [28] A. N. Startsev, On the distribution of the size of an epidemic in a non-Markovian model, *Theor. Prob. Appl.* **41**, 730 (1997).
- [29] G. Streftaris and G. J. Gibson, Non-exponential tolerance to infection in epidemic systems: Modeling, inference, and assessment, *Biostatistics* **13**, 580 (2012).
- [30] P. Van Mieghem and R. van de Bovenkamp, Non-Markovian Infection Spread Dramatically Alters the Susceptible-Infected-Susceptible Epidemic Threshold in Networks, *Phys. Rev. Lett.* **110**, 108701 (2013).
- [31] E. Cator, R. van de Bovenkamp, and P. Van Mieghem, Susceptible-Infected-Susceptible epidemics on networks with general infection and cure times, *Phys. Rev. E* **87**, 062816 (2013).
- [32] R. van de Bovenkamp, Epidemic processes on complex networks: Modelling, simulations and algorithms, Ph.D. thesis, Delft University of Technology, 2015.