Susceptible-infected-susceptible epidemics on networks with general infection and cure times

E. Cator,[1,*] R. van de Bovenkamp,[2,†] and P. Van Mieghem[2,‡]

[1]*Faculty of Science, Radboud University Nijmegen, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands*

[2]*Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands*

(Received 21 February 2013; revised manuscript received 29 March 2013; published 24 June 2013)

The classical, continuous-time susceptible-infected-susceptible (SIS) Markov epidemic model on an arbitrary network is extended to incorporate infection and curing or recovery times each characterized by a general distribution (rather than an exponential distribution as in Markov processes). This extension, called the generalized SIS (GSIS) model, is believed to have a much larger applicability to real-world epidemics (such as information spread in online social networks, real diseases, malware spread in computer networks, etc.) that likely do not feature exponential times. While the exact governing equations for the GSIS model are difficult to deduce due to their non-Markovian nature, accurate mean-field equations are derived that resemble our previous N-intertwined mean-field approximation (NIMFA) and so allow us to transfer the whole analytic machinery of the NIMFA to the GSIS model. In particular, we establish the criterion to compute the epidemic threshold in the GSIS model. Moreover, we show that the average number of infection attempts during a recovery time is the more natural key parameter, instead of the effective infection rate in the classical, continuous-time SIS Markov model. The relative simplicity of our mean-field results enables us to treat more general types of SIS epidemics, while offering an easier key parameter to measure the average activity of those general viral agents.

DOI: 10.1103/PhysRevE.87.062816 PACS number(s): 89.75.Hc, 05.40.−a, 87.19.X−

I. INTRODUCTION

Since the first modeling of epidemics on *general* complex networks by Pastor-Satorras and Vespignani [1], much effort has been devoted to better understanding the interplay between a simple dynamic process and the underlying network on which the process operates. Although the susceptible-infected-susceptible (SIS) model, which is one of the basis models in epidemiology (see, e.g., [2–5]), is only a first-order description of the spread of real epidemics, its simplicity and broad applicability justify these thorough analyses. Indeed, the range of applications is broad, ranging from biological virus spread [6] and information spread in social networks [7] to malware spread as an important threat for network security. While even the simple SIS model on an arbitrary network leads to an intractable analysis (due to state space explosion), mean-field approximations have been reasonably valuable in some networks. Here, we further demonstrate the power of mean-field theory in an important, practical generalization of the classical, continuous-time (Markovian) SIS model on networks.

In previous papers [8,9], the SIS model is described using exponential waiting times between infecting events and curing, recovery, or healing events. This ensures that the state of the process, describing for each node whether it is infected or not, is a Markov chain. We know, however, that these exponential distributions in general do not describe real-life epidemics well [10–12]. Here, we extend our results to general waiting times. Specifically, we generalize our previous N-intertwined mean-field approximation (NIMFA) and focus on the metastable state of the SIS epidemic process. The continuous-time SIS model on any network, in which the infection time T and the curing or recovery time R have a general distribution, is called the generalized SIS or GSIS model, of which the classical continuous-time SIS Markov model is a special case.

II. THE GSIS MODEL WITH GENERAL WAITING TIMES

Consider a graph G with N nodes and adjacency matrix A. Each node can be either infected or healthy; this is described by the vector X, such that $X_i = 1$ if node i is infected, and $X_i = 0$ if node i is healthy. If at time t node i gets infected, we draw independently of everything else a recovery time $R_i(t)$ and given $R_i(t)$, we independently draw, for each neighboring edge j, a random number $M_{ij}(t)$ of infecting times $T_{ij}^{(1)}(t) \leqslant \cdots \leqslant T_{ij}^{(M_{ij}(t))}(t) \leqslant R_i(t)$ such that at times $t + T_{ij}^{(k)}$ node i tries to infect node j. If node j is already infected at this time, nothing happens. Finally, at time $t + R_i(t)$ node i recovers and becomes healthy, but again susceptible to infection. We will consider the random vector of node i,

$$Z_i(t) = \left(R_i(t), Y_{ij_1}(t), \ldots, Y_{ij_{d_i}}(t)\right),$$

where d_i denotes the degree of node i, and hence the number of its neighbors, and where

$$Y_{ij}(t) = \left(T_{ij}^{(1)}(t), \ldots, T_{ij}^{(M_{ij})}(t)\right).$$

In particular, we confine our attention to the case where the distribution of $Z_i(t)$ does not depend on t, and if $t_1, \ldots, t_k$ are the times of infection for node i, we will assume $Z_i(t_1), \ldots, Z_i(t_k)$ to be independent and identically distributed (i.i.d.). Furthermore, we will assume that $Y_{ij_1}(t), \ldots, Y_{ij_{d_i}}(t)$ are i.i.d.

An exact analysis of the GSIS model on any network is very likely intractable, so that only an approximate treatment seems possible.

*e.cator@science.ru.nl
†r.vandebovenkamp@tudelft.nl
‡p. f. a.vanmieghem@tudelft.nl

III. MEAN-FIELD APPROXIMATION

We will determine the conditions for the distribution of Z_i such that a metastable state exists. Assuming that the metastable state exists, a kind of ergodicity of the process is required, for which it is relevant that the average recovery time of a node i is finite, $E[R_i] < \infty$. Furthermore, we need to control the number of infection events of a node i as well, so we assume that $E[M_{ij}] < \infty$. We denote by v_i the probability that node i is infected in the metastable state. The mean-field approximation [13,14] in this setting entails that, when we determine the effect of the neighbors on node i, we assume that node i does not influence its neighboring nodes in some relevant way. We will now explain what this assumption implies.

Consider node j as a neighbor of node i. In a large time interval $[0,S]$, the number of times node j was infected is asymptotically linear in S by the elementary renewal theorem [[15], p. 145]. Since the length of an infected period equals $E[R]$ (we omit the subscript for the node j, since the expectation is the same for each node), the number of infected periods is equal to $v_j S/E[R]$. During each infected period, node j will try to infect node i an average number $E[M]$ of times. This means that the total number of infection attempts from node j to node i asymptotically equals $v_j SE[M]/E[R]$. Now we apply the mean-field approximation: the fraction of infection events from j to i that is successful (i.e., node i was in fact healthy at the time of infection) equals $1 - v_i$. Hence, the total number of successful infections that node i will receive in the time interval $[0,S]$ will asymptotically be equal to

$$S \sum_{j \in U_i} \frac{E[M]}{E[R]} v_j (1 - v_i),$$

where U_i denotes the index set of neighbors of node i, so $U_i = \{j \mid a_{ij} = 1\}$. In the metastable state (where equilibrium holds), this number must equal the number of infected periods of node i in $[0,S]$, so that

$$S \sum_{j=1}^{N} a_{ij} \frac{E[M]}{E[R]} v_j (1 - v_i) = v_i \frac{S}{E[R]}.$$

In other words, we arrive, for any node $i \in G$, at

$$E[M](1 - v_i) \sum_{j=1}^{N} a_{ij} v_j = v_i, \tag{1}$$

which is exactly the same equation as in the N-intertwined mean-field approximation in the exponential case [16], if we replace $\tau = \beta/\delta$ by $E[M]$. The expected number of infection events in a Poisson process with intensity β within an exponential recovery time with expectation $1/\delta$ indeed equals the effective infection rate $\tau = \beta/\delta$.

The analogy with the NIMFA equations in [8,16] allows us to transfer the NIMFA analytic framework to the generalized SIS model. It follows from (1) and the epidemic threshold theorem in [8] that a *lower bound* for the epidemic threshold in GSIS epidemics satisfies

$$m_c = E[M_c] = \frac{1}{\lambda_1}, \tag{2}$$

where λ_1 is the largest eigenvalue of the adjacency matrix A of the graph G. Thus, if $E[M_c] > m_c$, then the epidemic process is eventually endemic (in the mean-field approximation), in which a nonzero fraction of the nodes remains infected, or the epidemic process dies out, after which the network is overall healthy. While $E[M_c]$ describes the activity of the viral agents in SIS epidemics, the right-hand side in the (mean-field) epidemic threshold equation (2) reflects the structural properties of the underlying network and emphasizes the role of the spectral radius λ_1 of the graph G.

In the remainder, we will assess the accuracy of the GSIS mean-field equations (1) and express the average number $E[M]$ of infection events during a healthy period R in terms of the probability distribution of the infection time T and the recovery time R. The main result is the general integral (4) for $E[M]$. When assuming a Weibullean infection time T and an exponential recovery time R, we will show in Sec. IV B1 that the analytic solution of the epidemic threshold equation (2) leads to the epidemic threshold scaling law (11) for large N that was earlier observed in [12] via extensive simulations.

IV. DETERMINATION OF $E[M]$

We wish to emphasize that we arrived at (1) without any assumption on the random vector Z, except for the existence of $E[M]$ and $E[R]$. However, it might be natural to consider a renewal process $T_1, T_2, \ldots$ starting at the time of infection, independent of the recovery time R, and running until $T_n \leqslant R < T_{n+1}$. In that case, we would have $M = n$. We will apply renewal theory [15, Chap. 8] to compute the number of infection events M in the random time R. The waiting time $W_n = \sum_{k=1}^{n} T_k$ for $n \geqslant 1$ is related to the counting process $\{M(t), t \geqslant 0\}$ by the renewal equivalence $\{M(t) \geqslant n\} \Longleftrightarrow \{W_n \leqslant t\}$. Indeed, the number $M(t)$ of infection events up to time t is at least n if and only if the nth renewal occurred at or before time t. By the law of total probability, we condition on the random recovery time R,

$$\Pr[W_n \leqslant R] = \int_0^\infty \Pr[W_n \leqslant u | R = u] \frac{d \Pr[R \leqslant u]}{du} du.$$

Since the time W_n of the nth infection attempt from node j to node i and the recovery time R of node i are independent, we have

$$\Pr[W_n \leqslant R] = \int_0^\infty \Pr[W_n \leqslant u] f_R(u) du,$$

where $f_R(u)$ denotes the probability density function of the recovery time R. Using the renewal equivalence $\{M(t) \geqslant n\} \Longleftrightarrow \{W_n \leqslant t\}$ and the expression for the mean in terms of the tail probabilities yields

$$\begin{aligned} E[M(R)] &= \sum_{k=1}^{\infty} \Pr[M(R) \geqslant k] = \sum_{k=1}^{\infty} \Pr[W_k \leqslant R] \\ &= \int_0^\infty \sum_{k=1}^{\infty} \Pr[W_k \leqslant u] f_R(u) du. \end{aligned} \tag{3}$$

If $f_T(u)$ is the probability density function of the infection time T and $\varphi_T(z) = E[e^{-zT}]$ the corresponding probability

generating function (Laplace transform), then [15, p. 138]

$$\Pr[W_k \leqslant u] = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{\varphi_T^k(z)}{z} e^{zu} dz,$$

where $c > 0$ and

$$\sum_{k=1}^{\infty} \Pr[W_k \leqslant u] = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{\varphi_T(z)}{1-\varphi_T(z)} \frac{e^{zu}}{z} dz.$$

Writing $E[M] = E[M(R)]$ and substituting the above into (3) yields

$$E[M] = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{\varphi_T(z)}{1-\varphi_T(z)} \frac{dz}{z} \int_0^{\infty} e^{zu} f_R(u) du.$$

Finally, we arrive at the general expression of the average number of infection attempts during a healthy, but susceptible period,

$$E[M] = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{\varphi_T(z)\varphi_R(-z)}{1-\varphi_T(z)} \frac{dz}{z}, \tag{4}$$

where $\varphi_R(z) = \int_0^{\infty} e^{-zu} f_R(u) du$ is the probability generating function (PGF) of the recovery time R.

Since any PGF obeys $|\varphi_T(z)| < 1$ for $\mathrm{Re}(z) > 0$, the integrand $g(z) = \frac{\varphi_T(z)\varphi_R(-z)}{z[1-\varphi_T(z)]}$ is analytic in the positive $\mathrm{Re}(z)$ plane, with the possible exception of the singularities of $\varphi_R(-z)$. Excluding the case of essential singularities or branch cuts of $\varphi_R(-z)$, we can deform the line of integration into a contour over the positive $\mathrm{Re}(z)$ plane [because the integrand $g(z) = g(re^{i\theta})$ vanishes for $r \to \infty$ at all angles $-\frac{\pi}{2} \leqslant \theta \leqslant \frac{\pi}{2}$]. Thus, (4) equals

$$E[M] = \frac{1}{2\pi i}\int_{C} \frac{\varphi_T(z)\varphi_R(-z)}{1-\varphi_T(z)} \frac{dz}{z} \tag{5}$$

where the contour C encloses the whole $\mathrm{Re}(z) > 0$ plane.

We now concentrate on some special cases in which the integral (4) or (5) for $E[M]$ can be evaluated.

A. The infection time T is exponential

If the infection time T is exponentially distributed with mean $\frac{1}{\beta}$, then the PGF equals $\varphi_T(z) = \frac{\beta}{z+\beta}$ and (4) simplifies to

$$E[M] = \frac{\beta}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{\varphi_R(-z)}{z^2} dz \quad (c > 0). \tag{6}$$

When we close the contour over the negative $\mathrm{Re}(z)$ plane, in which $\varphi_R(-z)$ is analytic and bounded $|\varphi_R(-z)| \leqslant 1$, we find, by Cauchy's integral theorem [17], that

$$E[M] = \beta \left.\frac{d\varphi_R(-z)}{dz}\right|_{z=0} = \beta E[R] = \tau$$

because $\varphi_R(-z) = E[e^{zR}]$ and $E[R] = \frac{1}{\delta}$. The result $E[M] = \tau$ is intuitively clear. Since the infection time T is exponentially distributed, the infection process is a Poisson process and the mean number of Poisson events in an interval equals its rate β multiplied by the average length of that interval, which is $E[R]$. Hence, if the infection time T is exponential or the infection follows a Poisson process, then, for any distribution of the recovery time R, we find that the NIMFA applies.

B. The recovery time R is exponential

If the recovery time R is exponentially distributed with mean $\frac{1}{\delta}$, then the PGF equals $\varphi_R(z) = \frac{\delta}{z+\delta}$ and (5) simplifies to

$$E[M] = \frac{\delta}{2\pi i}\int_{C} \frac{\varphi_T(z)}{z[1-\varphi_T(z)]} \frac{dz}{\delta - z}.$$

By Cauchy's residue theorem, we find that

$$E[M] = \frac{\varphi_T(\delta)}{1-\varphi_T(\delta)}, \tag{7}$$

from which a lower bound for the epidemic threshold follows from (2) by solving the equation

$$\varphi_T(\delta) = \frac{1}{1+\lambda_1}. \tag{8}$$

The relation (7) in the case of an exponential recovery time R with rate δ can be derived probabilistically, without resorting to contour integration. At the start of the infection, two cases can happen: either $T_1 < R$ or $T_1 > R$. In the latter case, $M = 0$. In the first case, $M \geqslant 1$ and at time T_1 everything starts anew: the distribution of $(R - T_1) \mid R > T_1$ is again exponential because of the memoryless property [15] of the exponential distribution. This means that M has a geometric distribution,

$$\Pr[M = k] = p^k(1-p) \quad \text{with} \quad p = \Pr[R > T_1] \quad \text{and} \quad k \geqslant 0.$$

Since all infection times $T_1, T_2, \ldots$ are i.i.d. and have the same distribution as T, we find the mean of a geometric random variable as

$$E[M] = \frac{\Pr[R > T]}{\Pr[R \leqslant T]} = \frac{\Pr[R > T]}{1 - \Pr[R > T]}.$$

This immediately shows that if T has an exponential distribution with parameter β, we find

$$E[M] = \frac{\beta}{\delta} = \tau.$$

Furthermore, for a general T, we have a nice connection between these probabilities and the Laplace transform:

$$\begin{aligned}\Pr[R > T] &= \int_0^{\infty} \Pr[R > s] f_T(s)\, ds \\ &= \int_0^{\infty} e^{-\delta s} f_T(s)\, ds \\ &= \varphi_T(\delta),\end{aligned}$$

from which (7) follows again.

1. The infection time T has a Weibull distribution

We will deduce an analytic law for the epidemic threshold, which was proposed in [12] based on simulations. While the curing process is still Poissonian with rate δ, the infection process at each node infects direct neighbors in a time T that is Weibullean [15, p. 56], with probability density function

$$f_T(x) = \frac{\alpha}{b}\left(\frac{x}{b}\right)^{\alpha-1} e^{-(x/b)^{\alpha}} \tag{9}$$

and mean $E[T] = b\Gamma(1 + \frac{1}{\alpha})$. For $\alpha = 1$, the Weibull distribution reduces to the exponential distribution. In order to

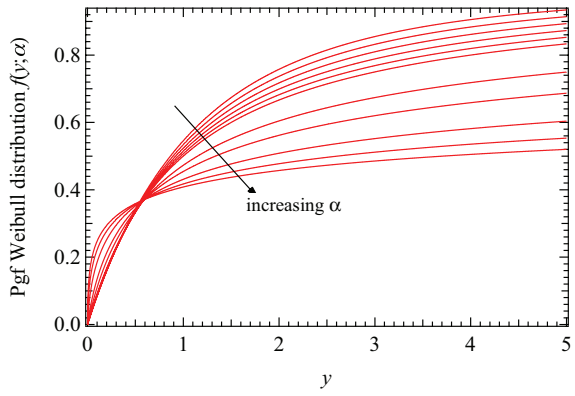

FIG. 1. (Color online) The PGF of the Weibull distribution as a function of y for various values of $\alpha = \{0.5,0.6,0.7,0.8,0.9,1,1.5,2,3,4,5\}$.

compare the Weibull with the exponential distribution, we fix the average infection time $E[T]$ to $\frac{1}{\beta}$, so that

$$b = \left[\Gamma\left(1+\frac{1}{\alpha}\right)\beta\right]^{-1}.$$

Thus, the parameter α in (9) tunes the power-law start and the tail of the Weibull distributions that all have the same mean infection time $E[T] = \frac{1}{\beta}$. Clearly, small α correspond to heavy tails (and large variance Var$[T]$), while large α correspond to almost deterministic infection times (with Var$[T] \to 0$).

The PGF of (9) is

$$\varphi_T(z) = \frac{\alpha}{b}\int_0^\infty e^{-zu}\left(\frac{u}{b}\right)^{\alpha-1} e^{-(u/b)^\alpha} du$$
$$= \alpha\int_0^\infty e^{-zbx-x^\alpha} x^{\alpha-1} dx$$

and

$$\varphi_T(\delta) = \alpha\int_0^\infty e^{-x/\Gamma(1+1/\alpha)\tau - x^\alpha} x^{\alpha-1} dx.$$

After substituting $u = x^\alpha$, we find with $y = \Gamma[1+\frac{1}{\alpha})\tau]^\alpha$ and for all α that

$$\varphi_T(\delta) = \int_0^\infty e^{-u^{1/\alpha}/\Gamma(1+1/\alpha)\tau - u} du$$
$$= y\int_0^\infty e^{-yx-x^{1/\alpha}} dx \triangleq f(y;\alpha), \qquad (10)$$

whose behavior is plotted in Fig. 1. The limiting case $\lim_{\alpha\to\infty} f(y;\alpha) = e^{-1}$, corresponding to a deterministic infection time $T = \frac{1}{\beta}$, explains the apparent intersection "point" in Fig. 1.

The epidemic threshold, expressed in units of $\tau = \frac{E[R]}{E[T]}$, follows from (8). We assume that $\frac{1}{1+\lambda_1}$ is small for sufficiently large N (and $\lambda_1 \gg 1$), which excludes lattices and any other graph whose spectral radius λ_1 does not increase with N. Since $f(y;\alpha)$ is monotonically increasing[1] in y from $f(0;\alpha) = 0$ towards $f(y;\alpha) = 1$ for $y \to \infty$, the assumption of small $\frac{1}{1+\lambda_1}$ requires the investigation of $f(y;\alpha)$ around $y = 0$. For small y and all α, we have

$$f(y;\alpha) = y\int_0^\infty e^{-x^{1/\alpha}}[1 - yx + O(y^2)]dx$$
$$= y\Gamma(\alpha+1) - \alpha\Gamma(2\alpha)y^2 + O(y^3).$$

The inversion $y = f^{-1}(\frac{1}{1+\lambda_1};\alpha)$ of (8) is to first order equal to

$$y \approx \frac{1}{\Gamma(\alpha+1)(1+\lambda_1)} \approx \frac{1}{\Gamma(\alpha+1)\lambda_1}.$$

Using the definition $y = [\Gamma(1+\frac{1}{\alpha})\tau]^\alpha$, we arrive, for large N and any graph, at the first-order *mean-field* epidemic threshold (in units of $\tau = \frac{E[R]}{E[T]}$) in a SIS process with Weibullean infection times T with mean $\frac{1}{\beta}$ and exponential recovery times R,

$$\tau_c^{(1)} = \frac{1}{\Gamma\left(1+\frac{1}{\alpha}\right)[\Gamma(\alpha+1)]^{1/\alpha}} \frac{1}{\lambda_1^{1/\alpha}}, \qquad (11)$$

where the superscript (1) refers to the first-order NIMFA mean-field approximation $\tau_c^{(1)}$ (see [14]). When $\alpha = 1$, we find again the classical mean-field epidemic threshold in a Markovian SIS process, $\tau_c^{(1)} = \frac{1}{\lambda_1} \leqslant \tau_c$, where τ_c is the exact SIS epidemic threshold. This analytic scaling law (11) supports our proposed fit in [12],

$$\tau_c \approx \frac{q(\alpha)}{\lambda_1^{1/\alpha}}$$

where $q(\alpha) = O(1)$. We mention that the prefactor $\frac{1}{\Gamma(1+1/\alpha)[\Gamma(\alpha+1)]^{1/\alpha}}$ is asymmetrically bell shaped in α with maximum, equal to 1, at $\alpha = 1$ and zero at $\alpha = 0$ and $\alpha \to \infty$, but slowly decaying as $\frac{e}{\alpha}$ towards zero for large α. In particular, the heavy-tailed regime ($\alpha < 1$) is of practical interest [11].

If the spectral radius λ_1 is not large (or the graph's size N is relatively small), then the inversion $y = f^{-1}(\frac{1}{1+\lambda_1};\alpha)$ needs to be computed numerically; most likely, in view of Fig. 1, the scaling (11) will not be observed. For example, for a complete graph K_{500} with $\lambda_1 = 499$, the epidemic threshold in terms of τ is shown in Fig. 2, where the line is computed as $\tau_c^{(1)} = \frac{[f^{-1}(1/500;\alpha)]^{1/\alpha}}{\Gamma(1+1/\alpha)}$. The simulations of the epidemic threshold in

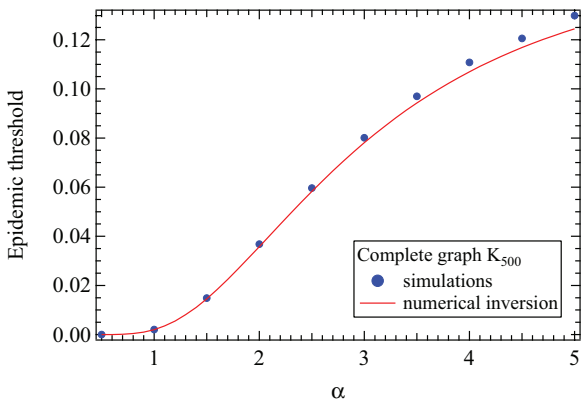

FIG. 2. (Color online) The epidemic threshold τ_c versus α, computed analytically as $\tau_c^{(1)} = [f^{-1}(1/500;\alpha)]^{1/\alpha}/\Gamma(1+1/\alpha)$ and via simulations.

[1]Recall that

$$f(y;\alpha) = \int_0^\infty e^{-(u/y)^{1/\alpha}-u} du$$

and the integrand is monotonically increasing in y for all α.

Fig. 2, deduced from Fig. 4, seem to indicate that the NIMFA lower-bounds the exact epidemic threshold for all values of $\alpha > 0$.

V. EVALUATION OF THE MEAN-FIELD EQUATION (1) AND $E[M]$

The mean-field equation (1) was compared with data generated by an event-based simulator. Every time a susceptible node becomes infected, an exponentially distributed curing time R is drawn and a curing event for the newly infected node is scheduled at that time. In addition to the curing time R, an infection time T is drawn from a Weibull distribution (9) for each neighbor. Only if the infection time T is smaller than the curing time R is the infection event scheduled. When the infection event is processed, a new infection time is drawn and, if the infection time is smaller than the curing time of the node that scheduled the infection event, a new infection event is scheduled.

To find the metastable fraction of infected nodes in the GSIS method, we log the exact percentage of the simulated time t_i that the network has spent in a state with i nodes infected and compute the average $\bar{y}(t)$ at time t as $\frac{1}{N}\sum_{j=0}^{N} j t_j$, the standard deviation of the fraction of infected nodes is calculated as $\frac{1}{N}\sqrt{\sum_{j=0}^{N} t_j[j-\bar{y}(t)]^2}$. We prevent the infection process from dying out by reinfecting the last remaining infected node immediately after it is cured; the elimination of the absorbing state of the GSIS process is called the modified GSIS process, as in [18]. To be certain that the process is indeed in the *steady state* of the modified GSIS process (corresponding to the *metastable state* of the GSIS process), we run two simultaneous but independent simulations on the same network and compare the time averages after an initial phase of 100 000 state changes. One process starts with 10% of the nodes infected, whereas the other process starts with all nodes infected. We conclude that the steady state is reached when the difference between the time averages of the two simulations is small, i.e., $\frac{|\bar{y}_1(t)-\bar{y}_2(t)|}{\bar{y}_1(t)+\bar{y}_2(t)} < 10^{-4}$. The steady-state fraction of infected nodes is then given by $\frac{\bar{y}_1+\bar{y}_2}{2}$. To avoid very long simulation times for larger values of α we set a time limit of 20 min simulation time per run.

Instead of evaluating the integral (4), the expected number of spreading events during the infection time of a node can be computed numerically as $E[M] = \sum_{k=0}^{\infty} k \Pr[M = k]$. As above, let M be the number of infection events over a link during the infection time R of a node, the waiting time until a curing event, and let T be the waiting time until the next spreading event. The probability that exactly k spreading events occur during an infected period follows (see [15, p. 140]) from the renewal equivalence $\{M(t) \geqslant n\} \Longleftrightarrow \{W_n \leqslant t\}$ as $\Pr[M = k] = \Pr[W_k \leqslant R] - \Pr[W_{k+1} \leqslant R]$, where $W_n = \sum_{k=1}^{n} T_k$ denotes the sum of n independent realizations of T. Let $f_Y(x) = \frac{d}{dx}\Pr[Y \leqslant x]$ denote the probability density function (PDF) of the random variable Y. Since the PDF of the sum of two independent random variables is given by the convolution of the two PDFs of the random variables, we can express the PDF of W_k as a series of convolutions, i.e., $f_{W_3}(x) = f_T(x) * f_T(x) * f_T(x)$. To specify $\Pr[W_k \leqslant R]$, we can use a similar approach. Let

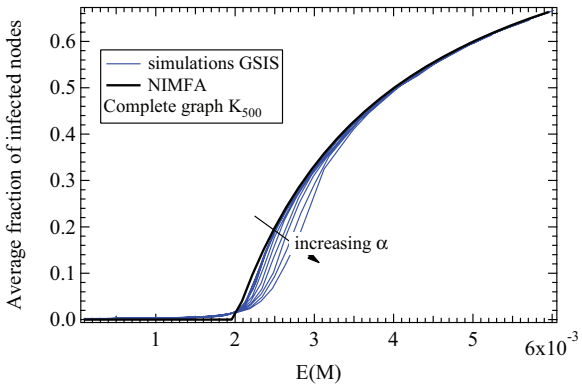

FIG. 3. (Color online) The average metastable state fraction of infected nodes versus $E[M]$ in the complete graph with $N = 500$ nodes.

$Y_k = W_k - R$; then $\Pr[W_k \leqslant R] = \Pr[Y_k \leqslant 0]$. Since the PDF of $-R$ is $f_R(-x)$ and $f_{Y_k}(x) = f_{W_k}(x) * f_R(-x)$, we find that $\Pr[W_k \leqslant R] = \int_{-\infty}^{0} f_{Y_k}(x)dx$. In general, the series of convolutions can be efficiently determined numerically using fast Fourier transforms. The number of terms in the sum needed to closely approximate $E[M] = \sum_{k=0}^{\infty} k \Pr[M = k]$ is typically on the order of 10.

Just as in the NIMFA, which was shown (see, e.g., [12]) to be an upper bound of the probability of infection v_i of node i, the mean-field equations (1) of the GSIS method are also upper-bounding the infection probability as illustrated in Figs. 3 and 4, which both contain the same data. Figure 3 emphasizes that $E[M]$ is the more natural parameter, instead of the ratio $\tau = \frac{E[R]}{E[T]}$, because all curves tend to each other, indicating that $E[M]$ constitutes a proper scaling.

When the infection time T tends to a deterministic time (i.e., $\alpha \to \infty$), the variance in the simulations increases so that more realizations need to be simulated for increasing α. Still, the mean of the simulation seems nicely upper bounded by the NIMFA result. While we have concentrated here on a comparison based on the complete graph [for which the NIMFA is exactly available [16] and the average metastable state fraction of infected nodes is $y_\infty(\tau) = \frac{1}{N}\sum_{j=1}^{N} v_{j\infty} = 1 - \frac{1}{(N-1)\tau}$], the same agreement is found on Erdős-Rényi graphs and Barabasi-Albert graphs as illustrated in Fig. 5.

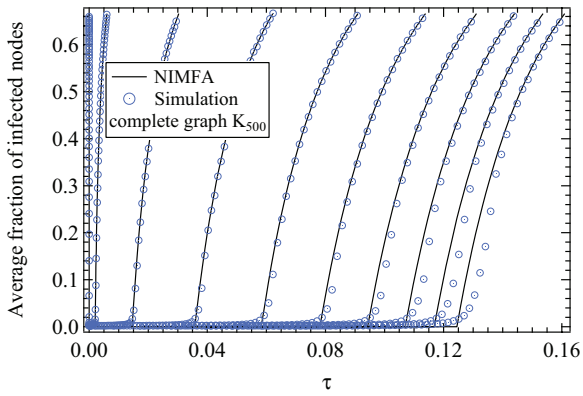

FIG. 4. (Color online) The average metastable fraction of infected nodes versus τ in the complete graph K_{500} for various α ranging from 0.5 (left) to 5 (right) with steps of 0.5.

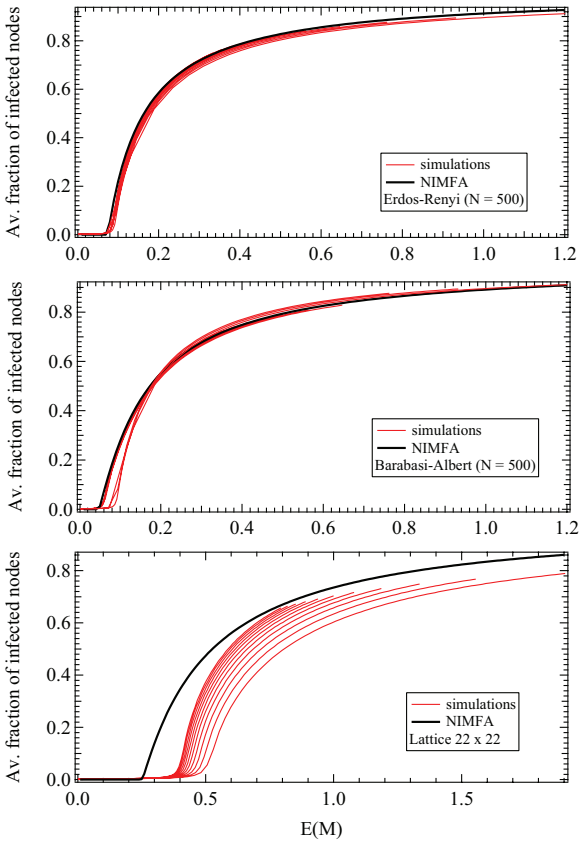

FIG. 5. (Color online) The average fraction of infected nodes versus $E[M]$ for three different graph types (with $N = 500$ and 484).

For lattices, on the other hand, Fig. 5 shows (as also found previously; see, e.g., [12]) that mean-field approximations are less accurate. Finally, GSIS epidemics on the complete graph K_N with two nonexponential distributions for T and R also follow the same behavior as shown in Fig. 3.

VI. CONCLUSION

While the classical, continuous-time SIS Markov model possesses both exponential infection and curing and recovery times, we have extended this SIS model to a generalized continuous-time SIS model in which the infection and curing and recovery times each have a general distribution. Many real-world epidemics (such as information spread on Twitter [11], real diseases, malware spread in computer networks, etc.) do not feature exponential times, which underlines the usefulness of a generalized SIS model on networks.

We have derived mean-field equations (1) that resemble our previous NIMFA equations, so that the whole analytic machinery of the NIMFA, surveyed in [16], is applicable. These mean-field equations are compared with precise simulations and show a remarkable agreement, similar to the accuracy of the NIMFA for exponential times. Although the analysis here has concentrated on the simplest epidemics, namely, SIS epidemics, the generalization of the NIMFA to multilayer and multicompartment epidemics in [19] as well as to nonhomogeneous epidemics [20] suggests that the current non-Markovian extensions will apply equally well to more general epidemics, mainly by replacing τ by $E[M]$ in the equations.

The average number $E[M]$ of infection attempts during a recovery time R is shown to be a key characterizing parameter in the GSIS model, which replaces the Poisson rates in the classical SIS model. In other words, for a general SIS epidemic process, $E[M]$ is the more natural parameter, instead of the ratio $\tau = \frac{E[R]}{E[T]}$, called the effective infection rate in the classical, continuous-time SIS Markov model. Only when the infection process is Poissonian with infection rate β and curing rate δ, as derived in Sec. IV A, does $E[M]$ equal the ratio $\tau = \frac{E[R]}{E[T]} = \frac{\beta}{\delta}$.

Another practical conclusion is that, irrespective of our knowledge of the underlying distributions for the infection and recovery times, we can determine the average metastable state fraction of infected nodes, based only on an estimate of the average number $E[M]$ of infection attempts during a recovery time R. In most practical cases, the quantity $E[M]$ is easier to measure than the ratio $\tau = \frac{E[R]}{E[T]}$. The generalized epidemic threshold rule (2) provides a useful criterion to verify whether a certain viral agent, whose precise properties are unknown except for its average activity $E[M]$, will cause a pandemic in a network or not.

We believe that the generalization of the classical SIS process towards the GSIS model is an important step forwards, while the main computational tools (in the mean-field approximation) are almost the same (just a replacement of $\tau = \frac{E[R]}{E[T]}$ by $E[M]$). In other words, the shape of the curve of the average metastable state fraction of infected nodes is mainly determined by the underlying topology and can be computed via the NIMFA, whereas the scaling via $E[M]$ reflects the epidemic details, which allow us to compare viral agents with different epidemic properties (measured via T and R) on the same contact network. Moreover, the exact governing equations for the non-Markovian GSIS process are believed to be intractable to solve. Consequently, the assessment of the accuracy of the NIMFA is still an important open problem for the simple SIS epidemic model on networks: on any topology on which the NIMFA is accurate for the SIS process, it seems accurate for the GSIS process as well. We would like to have a general mean-field criterion that specifies for which types of graphs the NIMFA is sufficiently accurate.

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

[2] D. J. Daley and J. Gani, *Epidemic Modeling: An Introduction* (Cambridge University Press, Cambridge, UK, 1999).

[3] N. T. J. Bailey, *The Mathematical Theory of Infectious Diseases and its Applications*, 2nd ed. (Charlin Griffin & Company, London, 1975).

[4] R. M. Anderson and R. M. May, *Infectious Diseases of Humans: Dynamics and Control* (Oxford University Press, Oxford, UK, 1991).

[5] O. Diekmann, H. Heesterbeek, and T. Britton, *Mathematical Tools for Understanding Infectious Disease Dynamics* (Princeton University Press, Princeton, NJ, 2012).

[6] S. Sutrave, C. Scoglio, S. A. Isard, J. M. Shawn Hutchinson, and K. A. Garrett, PLoS One **7**, e37793 (2012).
[7] J. L. Iribarren and E. Moro, Phys. Rev. E **84**, 04116 (2011).
[8] P. Van Mieghem, J. Omic, and R. E. Kooij, IEEE/ACM Trans. Netw. **17**, 1 (2009).
[9] P. Van Mieghem and E. Cator, Phys. Rev. E **86**, 016116 (2012).
[10] E. Limpert, W. A. Stahel, and M. Abbt, Bioscience **51**, 341 (2001).
[11] C. Doerr, N. Blenn, and P. Van Mieghem, PLoS One **8**, e64349 (2013).
[12] P. Van Mieghem and R. van de Bovenkamp, Phys. Rev. Lett. **110**, 108701 (2013).
[13] P. Van Mieghem, Europhys. Lett. **97**, 48004 (2012).
[14] E. Cator and P. Van Mieghem, Phys. Rev. E **85**, 056111 (2012).
[15] P. Van Mieghem, *Performance Analysis of Communications Systems and Networks* (Cambridge University Press, Cambridge, UK, 2006).
[16] P. Van Mieghem, Computing **93**, 147 (2011).
[17] E. C. Titchmarsh, *The Theory of Functions* (Oxford University Press, Amen House, London 1964).
[18] E. Cator and P. Van Mieghem, Phys. Rev. E **87**, 012811 (2013).
[19] F. D. Sahneh, C. Scoglio, and P. Van Mieghem, IEEE/ACM Trans. Netw. 1 (2013), doi: 10.1109/TNET.2013.2239658.
[20] P. Van Mieghem and J. Omic, Report No. 2008081, Delft University of Technology, 2008, arXiv:1306.2588, www.nas.ewi.tudelft.nl/people/Piet/TUDelftReports.