Approximate formula and bounds for the time-varying susceptible-infected-susceptible prevalence in networks

P. Van Mieghem*
Delft University of Technology, Faculty of EECS, P.O. Box 5031, 2600 GA Delft, The Netherlands
(Received 1 February 2016; revised manuscript received 14 April 2016; published 26 May 2016)

Based on a recent exact differential equation, the time dependence of the SIS prevalence, the average fraction of infected nodes, in any graph is first studied and then upper and lower bounded by an explicit analytic function of time. That new approximate "tanh formula" obeys a Riccati differential equation and bears resemblance to the classical expression in epidemiology of Kermack and McKendrick [Proc. R. Soc. London A **115**, 700 (1927)] but enhanced with graph specific properties, such as the algebraic connectivity, the second smallest eigenvalue of the Laplacian of the graph. We further revisit the challenge of finding tight upper bounds for the SIS (and SIR) epidemic threshold for all graphs. We propose two new upper bounds and show the importance of the variance of the number of infected nodes. Finally, a formula for the epidemic threshold in the cycle (or ring graph) is presented.

DOI: 10.1103/PhysRevE.93.052312

I. THE MARKOVIAN SIS PROCESS ON A GRAPH

Epidemic spreading processes on graphs are popular to model diffusion phenomena in real-world networks [1], ranging from biological virus infections, to information spread in social and communications networks, to cascading failures in infrastructural networks. However, even one of the simplest epidemic processes on a graph, the Susceptible-Infected-Susceptible (SIS) model defined below, is not completely understood. Mean-field approximations primarily dominate the analyses in papers, and often their approximate nature is ignored or forgotten. Yet, in some type of graphs such as the cycle [2] and the star [3], discrepancies between mean-field approximations and the Markovian SIS process can be significant. Starting from a recent exact result [4], we analyze here the average fraction of infected nodes of the Markovian SIS process on any graph, trying to understand for which graphs the mean-field approximation is accurate [5].

We consider an unweighted, undirected graph G containing a set $\mathcal{N}$ of N nodes (also called vertices) and a set $\mathcal{L}$ of L links (or edges). The topology of the graph G is represented by a symmetric $N \times N$ adjacency matrix A. In an SIS epidemic process [1,6–9] on the graph G, the viral state of a node i at time t is specified by a Bernoulli random variable $X_i(t) \in \{0,1\}$: $X_i(t) = 0$ for a healthy, but susceptible node and $X_i(t) = 1$ for an infected node. A node i at time t can be in one of the two states: *infected*, with probability $\Pr[X_i(t) = 1]$, or *healthy*, with probability $\Pr[X_i(t) = 0] = 1 - \Pr[X_i(t) = 1]$, but susceptible to the infection. We assume that the curing process per node i is a Poisson process with rate δ and that the infection rate per link is a Poisson process with rate β. Obviously, only an infected node can infect its healthy direct neighbors. Both the curing and infection Poisson process are independent. The effective infection rate, sometimes also called the spreading rate [1], is defined by $\tau = \frac{\beta}{\delta}$. This description defines the continuous-time, Markovian SIS epidemic process on a graph G. We do not consider non-Markovian epidemics [10,11].

One of the most intriguing features of the SIS process is the occurrence of an epidemic threshold at τ_c: for effective infection rates $\tau < \tau_c$, the infection dies out exponentially fast for sufficiently large times, whereas for $\tau > \tau_c$, the infection stays very long in any sufficiently large network [12,13] but is eventually wiped out in any finite graph due to the existence of an absorbing state (the overall healthy state). The regime of persistent infection ($\tau > \tau_c$), called the metastable or quasistationary state, is reached rapidly given an initial set of infected nodes [14] and can last extremely long in large networks. Only in the limit $N \to \infty$ is the epidemic phase transition sharp, precisely defining τ_c, and the infection can remain forever for $\tau > \tau_c$ with a nonzero probability (similar to a supercritical branching process [15, ch. 12]). For finite sizes of a network, the phase transition occurs in a τ region around τ_c, featuring a positive width that decreases with N and that complicates the precise definition of τ_c. In spite of the crucial importance of the knowledge of the epidemic threshold τ_c, for surprisingly few graphs is the scaling of τ_c with size N known [3]. The only solid general result is the mean-field lower bound, $\tau_c \geqslant \tau_c^{(1)} = \frac{1}{\lambda_1}$, where λ_1 is the spectral radius of the adjacency matrix A of the graph and $\tau_c^{(1)}$ refers to the first order N-intertwined mean-field approximation (NIMFA) [15,16]. A second order improvement $\tau_c^{(2)} \geqslant \tau_c^{(1)}$ is possible [17], though at the expense of much larger computations that provide less insight than the first order NIMFA. As mentioned in Ref. [4], the *lower* bound $\tau_c^{(1)} = \frac{1}{\lambda_1}$ is of great practical use: if the effective infection rate τ can be controlled such that $\tau \leqslant \tau_c^{(1)}$ or the network can be designed to lower the spectral radius λ_1 of its graph [18], then the network is safeguarded from long-term, massive infection. On the other hand, knowing that an infection will persist almost surely in the network and that the epidemic process is thus operating above the epidemic threshold can be of equal importance. The challenge to determine a tight, but general upper bound for τ_c (in any graph) is addressed here.

Explicitly using the properties of the Bernoulli random variable $X_i(t)$, for which the rather difficult probability operator $\Pr[\cdot]$ can be replaced by the linear expectation operator $E[\cdot]$ as illustrated earlier [4,17], the exact Markovian

*P.F.A.VanMieghem@tudelft.nl.

SIS-governing equation [15,17] for the infection probability of node i is

$$\frac{dE[X_i(t)]}{dt} = E\left[-\delta X_i(t) + \beta(1 - X_i(t))\sum_{k=1}^{N} a_{ki} X_k(t)\right]. \quad (1)$$

The time derivative of the infection probability $E[X_i(t)] = \Pr[X_i(t) = 1]$ of a node i consists of the expectation of two competing processes in (1), expressed in the Bernoulli random variable $X_i \in \{0,1\}$: (a) while node i is infected, i.e., $X_i(t)$, the node i is cured at rate δ and (b) while node i is healthy, i.e., not infected $(1 - X_i(t))$, all infected neighbors $\sum_{k=1}^{N} a_{ki} X_k(t)$ of node i try to infect the node i with rate β. We define the fraction of infected nodes by

$$S(t;\tau) = \frac{1}{N}\sum_{i=1}^{N} X_i(t), \quad (2)$$

where the dependence on the effective infection rate τ has been made explicit. We denote the average fraction of infected nodes, also called the *prevalence*, by

$$y(t;\tau) = E[S(t;\tau)] = \frac{1}{N}\sum_{i=1}^{N} \Pr[X_i(t) = 1]. \quad (3)$$

The prevalence obeys the differential equation

$$\frac{dy(t^*;\tau)}{dt^*} = -y(t^*;\tau) + \frac{\tau}{N} E[w(t^*;\tau)^T Q w(t^*;\tau)], \quad (4)$$

where $t^* = t\delta$ is the normalized time measured in units of the average occurring time δ^{-1}, $Q = \Delta - A$ is the Laplacian of the graph G with $\Delta = \mathrm{diag}(d_1, d_2, \ldots, d_N)$, d_i is the degree of node i in G, and the Bernoulli vector $w = (X_1, X_2, \ldots, X_N)$. The proof of (4) is included in Appendix A, while a slightly more general proof for both SIS and SIR is presented in Ref. [4]. The state $w(t;\tau)$ at time t depends on the initial condition, i.e., the state $w(0;\tau)$ at $t = 0$. In the sequel, the time and τ dependence as well as the dependence on the initial condition is sometimes omitted to simplify the notation. We denote x_k the kth eigenvector of the symmetric $N \times N$ Laplacian Q belonging to the eigenvalue μ_k, ordered as $\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_{N-1} \geqslant \mu_N = 0$. The eigenvectors are normalized and obey the orthogonality requirement $x_k^T x_m = \delta_{km}$, where δ_{km} is the Kronecker delta, which is equal to one, if $k = m$, and otherwise $\delta_{km} = 0$. The eigenvector belonging to the zero eigenvalue $\mu_N = 0$ equals $x_{N-1} = \frac{u}{\sqrt{N}}$, where $u = (1,1,\ldots,1)$ is the all-one vector. The second smallest eigenvalue μ_{N-1} of the Laplacian Q is called by Fiedler [19] the algebraic connectivity. The eigenvalues of the corresponding $N \times N$ adjacency matrix A are $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_N$.

Physically we can interpret the term $\frac{\tau}{N} E[w^T Q w]$ in (4) as the force per node that drives the infection. In particular, introducing the basic Laplacian property $z^T Q z = \sum_{l \in \mathcal{L}} (z_{l^+} - z_{l^-})^2$ for any real vector z, where the link l connects the node l^+ and node l^-, yields

$$E[w^T Q w] = 2\sum_{l \in \mathcal{L}} E[X_{l^+}(1 - X_{l^-})],$$

illustrating that only links with one end infected contribute to the average fraction of infected nodes. Thus, $w(t^*;\tau)^T Q w(t^*;\tau)$ equals twice the cut size, i.e., the number of links in the cut between infected and healthy nodes in the graph G at time t^*, so that the maximum increase of the prevalence, $\max_{t^*} \frac{dy(t^*;\tau)}{dt^*}$, is related to the maximum cut size, whose determination is an NP-complete problem.

We first derive some general considerations on the exact SIS prevalence differential equation (4) in Sec. II. In Sec. III we present a new analytic, but approximate, "tanh formula" (13) for the prevalence in any network in Sec. III, deduced from (4) and derived in Appendix C. The new approximation (13) of the prevalence has a surprisingly classical form, which already appeared in the pioneering paper by Kermack and McKendrick [20], showing the influence of the underlying graph on the prevalence, mainly via the algebraic connectivity μ_{N-1}, but also larger eigenvalues of the Laplacian Q. In Figs. 1 and 2, we illustrate the potential of this new formula, in comparison with the well-established mean-field approximation NIMFA. In the second part (Secs. IV and V), we propose new *upper* bounds of the exact SIS epidemic threshold τ_c. Some mathematical derivations are deferred to the appendices, and the theorems and lemmas are proved in Appendix D.

II. GOVERNING EQUATION (4) OF THE PREVALENCE

A. Regime $\tau < \tau_c$ below the epidemic threshold

Since $1 - X_i \leqslant 1$, the governing equation (1) is upper bounded by

$$\frac{dE[X_i]}{dt} \leqslant E\left[-\delta X_i + \beta\sum_{k=1}^{N} a_{ki} X_k\right],$$

from which the upper bound of the vector w follows as

$$\frac{d}{dt^*} E[w(t^*;\tau)] \leqslant (\tau A - I) E[w(t^*;\tau)].$$

The solution of this linear differential inequality for the vector of infection probabilities is

$$E[w(t^*;\tau)] \leqslant e^{(\tau A - I)t^*} E[w(0)]. \quad (5)$$

For a characteristic polynomial $\det(A - \lambda I) = \prod_{j=1}^{s} (\lambda - \lambda_j)^{m_j}$, where each eigenvalue λ_j is different with multiplicity m_j and $\sum_{j=1}^{s} m_j = N$, the exponential matrix function is [21, p. 116]

$$e^{At} = \sum_{j=1}^{s} \left\{\sum_{k=1}^{m_j} Z_{sk} t^{k-1}\right\} e^{\lambda_s t}, \quad (6)$$

where the matrices Z_{sk} are linearly independent constant matrices that are polynomials in A. Hence, as also reported earlier in Refs. [10] and [15, pp. 457–458], the epidemic dies out in any graph exponential fast in t^* as $O(e^{(\lambda_1 \tau - 1)t^*})$ for large t^* when $\tau < \frac{1}{\lambda_1}$. However, even below the epidemic threshold $\tau < \tau_c$, the infection probabilities $E[X_i(t)]$, and thus the prevalence $y(t)$ also can initially (for small $t > 0$) increase due to the factors t^{k-1} in (6) when the adjacency matrix A has eigenvalues with multiplicity $m_j > 1$. An increase in the prevalence $y(t)$ below the epidemic threshold occurs, for example, in the star graph $K_{1,N-1}$, where the zero eigenvalue of the adjacency matrix A has a multiplicity of $N - 2$.

The spectral counterpart (B15) of the SIS prevalence governing equation (4) shows that, if $\tau\mu_k - 1 < 0$ for all positive Laplacian eigenvalues $1 \leqslant k \leqslant N-1$, thus if $\tau < \frac{1}{\mu_1}$, then $\frac{dy(t^*;\tau)}{dt^*} \leqslant 0$ and the prevalence $y(t^*;\tau)$ is nonincreasing for all times t^*. Since $N \geqslant \mu_1 > d_{\max} \geqslant \lambda_1$, a decrease in the prevalence at any time corresponds with an operation well below the epidemic threshold τ_c.

B. Properties of the prevalence y in a connected graph

Multiplying both sides in (4) by e^{t^*} yields, after rearrangement,

$$\frac{d}{dt^*}\{e^{t^*}y(t^*;\tau)\} = e^{t^*}\frac{\tau}{N}E[w(t^*;\tau)^T Q w(t^*;\tau)].$$

Integrating both sides over $[0,t^*]$ leads to

$$y(t^*;\tau) = e^{-t^*}y(0;\tau) + \frac{\tau}{N}\int_0^{t^*} e^{-(t^*-u)}E[w(u;\tau)^T Q w(u;\tau)]\,du. \quad (7)$$

The integral in (7) is a convolution and seems similar to the general integral representation of epidemic processes [9]. Although the first term in (7) decreases exponentially fast, we cannot conclude that the influence of the initial condition disappears as fast as $O(e^{-t^*})$, because also $E[w(u;\tau)^T Q w(u;\tau)]$ depends on $w(0)$ in a unknown way. The quadratic non-negative form $w^T Q w \geqslant 0$, as illustrated in (B9) in Appendix B, implies that $e^{t^*}y(t^*;\tau)$ is a nondecreasing function of the time t^* with only extrema at times when $E[w^T Q w] = 0$. The average of a non-negative random variable is only zero if the random variable itself is zero. Now, $w^T Q w = 0$ only if the SIS Markovian process is in either of two states: the absorbing state and the all-infected state. Only if the process starts in the absorbing state (a trivial process in which there is no infection), there holds that $w^T Q w = 0$ for any time $t > 0$ and $E[w^T Q w] = 0$. In all other cases, the process cannot remain in a single state so that the average (over all possible realizations) at a finite time t is $E[w^T Q w] > 0$: even when each realization of the process starts at a same state, not all realizations will reach another same state at a time t. In particular, not all realizations reach the all-infected state or the absorbing state at the same time. Since $E[w^T Q w] > 0$ [excluding absence of infection, for which $y(t^*;\tau) = 0$ for all $t^* \geqslant 0$], we conclude from (7) that

$$y(t^*;\tau) > e^{-t^*}y(0;\tau),$$

so that the prevalence cannot decrease faster than $O(e^{-t^*})$, limiting the extinction speed of the epidemic in any network. The lower bound corresponds to the $\tau = 0$ regime in (5). Since the SIS process has an absorbing state, each realization in a finite graph will eventually hit the absorbing state. Since the support of the absorption time in a graph is $(0,\infty)$, because a realization can die out fast while another can remain oscillating between certain states indefinitely long, we have that

$$\lim_{t^*\to\infty} y(t^*;\tau) = 0,$$

consistent with the above inequality that tells us that the prevalence is always positive for finite times.

Let us focus on extremal points at time $t^* = \theta$, for which $\frac{dy(t^*;\tau)}{dt^*}|_{t^*=\theta} = 0$. By (4), each extremal time point θ satisfies

$$y(\theta;\tau) = \frac{\tau}{N}E[w(\theta;\tau)^T Q w(\theta;\tau)]. \quad (8)$$

The previous arguments show that $\theta \to \infty$ is an extremal point. The condition (8) equates at time θ the averages of the total healing rate $\delta N y(\theta;\tau)$ (i.e., the total number of infected nodes that cures per unit time) and the total infection rate $\beta E[w(\theta;\tau)^T Q w(\theta;\tau)]$ in the graph G and reflects a temporal balance between the cut size and the set of infected nodes. The quadratic form (B14) indicates for $\tau > 0$ that

$$E\left[w(\theta;\tau)^T\left(Q - \frac{1}{\tau}I\right)w(\theta;\tau)\right] = 0,$$

which may seem to suggest that $w(\theta;\tau)$ is "close" to the eigenvector of Q belonging to the eigenvalue close to τ^{-1}. However, the resemblance to the eigenvalue equation is misleading: w is a binary vector, whereas the normalized eigenvectors of Q possess components with real values in $[-1,1]$, with at least one negative component (except for the eigenvector $x_N = u$ belonging to eigenvalue $\mu_N = 0$). Defining the binary matrix $W = ww^T$ and introducing the Hadamard matrix product, we can write

$$w(\theta;\tau)^T\left(Q - \frac{1}{\tau}I\right)w(\theta;\tau) = \sum_{i=1}^{N}\sum_{j=1}^{N} X_i X_j\left(Q - \frac{1}{\tau}I\right)_{ij} = u^T\left[W \circ \left(Q - \frac{1}{\tau}I\right)\right]u,$$

which sums all elements in the filtered matrix $W \circ (Q - \frac{1}{\tau}I)$, which is a submatrix of $(Q - \frac{1}{\tau}I)$, corresponding to the subgraph of G containing all infected nodes.

Lemma 1. Let the scalar product of the infection state vector $w(t^*)$ at time t^* and the Laplacian eigenvector x_k belonging to the k-largest eigenvalue μ_k be $\zeta_k(t^*) = w^T(t^*)x_k$. At an extremal time point θ, the prevalence is

$$y(\theta;\tau) = \frac{1}{2}\left(1 - \frac{1}{\tau\mu_{N-1}}\right) + \frac{1}{2}\sqrt{\left(1 - \frac{1}{\tau\mu_{N-1}}\right)^2 - \frac{4\Psi(\theta)}{\tau\mu_{N-1}}}, \quad (9)$$

where

$$\Psi(t^*) = \tau\mu_{N-1}\left\{\mathrm{var}[S(t^*)] - \frac{E[R(t^*)]}{N\mu_{N-1}}\right\} \quad (10)$$

and the correction R equals

$$R(t^*) = \sum_{k=1}^{N-2}(\mu_k - \mu_{N-1})\zeta_k^2(t^*). \quad (11)$$

Lemma 1, proved in Appendix D 1, specifies $y(\theta;\tau)$ in terms of a new spectral function Ψ. If there is a maximum $y(\theta;\tau) \geqslant 0$ for $\tau < \frac{1}{\mu_{N-1}}$, then Lemma 1 indicates that $\Psi(\theta) < 0$, whereas a maximum $y(\theta;\tau)$ for $\tau > \frac{1}{\mu_{N-1}}$ implies that $\Psi(\theta) > 0$. Thus, in general, $\Psi(t^*)$ can be positive as well as negative. As shown below in Sec. III, the function Ψ plays an important role in the assessment of our new approximate formula for the prevalence. Moreover, the function Ψ strongly depends on the variance

var[S], another quantity that will appear later in upper bounds for the epidemic threshold (see, e.g., Theorem 1).

For $\tau > \tau_c$, let the time t_m^* mark the beginning of the metastable or quasistationary regime, which can last (very) long [13]. In general, the prevalence at t_m^* is not necessarily a maximum, because it depends on the initial condition $y(0;\tau) = y_0$: if $y_0 > y(t_m^*)$, then $y(t^*)$ may be decreasing for all $t^* > 0$, whereas, if $y_0 < y(t_m^*)$, then a maximum in $y(t^*)$ is reached at $t^* = t_m^*$. In addition, if t_m^* corresponds to the single maximum, then $\frac{dy(t^*;\tau)}{dt^*} < 0$ for all $t^* > t_m^*$: during the metastable regime, the prevalence is strictly decreasing, although (very) slowly [13], to zero (at $t \to \infty$). The governing equation (4) then shows that, for all $t^* > t_m^*$,

$$y(t^*;\tau) \geqslant \frac{\tau}{N} E[w(t^*)^T Q w(t^*)],$$

illustrating that lower bounding the prevalence y is possible in any graph for $\tau > \tau_c$.

An interesting open question is "given an initial condition $w(0)$ and a finite graph G, how many extremal points obeying (8) has the SIS process?" We believe (but cannot prove) that, in a SIS epidemic process on any connected graph G with fixed τ, there are no multiple extrema (i.e., more than two extremal points) for a positive prevalence y. We remark that the situation when varying the effective infection rate τ is different: $y(t^*;\tau)$ can have multiple extrema as a function of τ. Indeed, consider a graph consisting of two complete graphs K_m and K_n with $n > m$, connected by a path P_h with h hops or links. When $\frac{1}{n} \leqslant \tau_{c_{K_n}} < \tau < \tau_{c_{K_m}} \leqslant \frac{1}{m}$, the epidemic spreads in K_n, but not in K_m. Even for $\tau > \frac{1}{m}$ and initially only a node in K_n is infected, the epidemic may never reach K_m if the path P_h is long enough. Only if $\tau >> \tau_{c_{P_h}} \geqslant \frac{1}{2}$, then both cliques K_m and K_n remain infected for a long time.

III. NEW APPROXIMATE FORMULA FOR THE SIS PREVALENCE

The governing equation (4) of the prevalence depends on a quadratic form $w^T Q w$ of the Laplacian Q of a underlying graph G. Introducing spectral graph theory, which supplies many insights and properties of this quadratic form (see Appendix B and Refs. [22,23]), leads to a new formula of the SIS prevalence, which we derive here.

Introducing the expression (B12) for the quadratic form $w^T Q w$, derived in Appendix B, into the differential equation (4), with the definition (B13) of R in Lemma 1 and using $E[S - S^2] = y - E[S^2]$ and $E[S^2] = y^2 +$ var[S], yields

$$\frac{dy}{dt^*} = (\tau\mu_{N-1} - 1)y - \tau\mu_{N-1}y^2 - \Psi, \tag{12}$$

where the function Ψ, specified in (10), is generally a rather complicated function of t^* and $y(t^*)$, which we approximate below. Clearly, if $\frac{dy}{dt^*} = 0$, then $y(t^*)$ remains constant and (12) indicates that, then, $\Psi(t^*)$ also does not change with time. Hence, in the metastable state, where $\frac{dy}{dt^*} = \varepsilon$ and ε is very small and negative, the function $\Psi(t^*)$ hardly changes with time t^*.

Suppose that we can bound

$$c_L \leqslant \Psi \leqslant c_U,$$

where c_L and c_U are constants, then the prevalence $y(t^*)$ can be bounded, for the same initial condition y_0, by

$$\tilde{y}(t^*) = \frac{1}{2}\left(1 - \frac{1}{\tau\mu_{N-1}}\right) + \frac{\Xi}{2}\tanh\left\{\frac{\tau\mu_{N-1}\Xi}{2}t^* + \operatorname{arctanh}\left[\frac{2y_0 - \left(1 - \frac{1}{\tau\mu_{N-1}}\right)}{\Xi}\right]\right\} \tag{13}$$

for $c = c_L$ and $c = c_U$, where

$$\Xi = \sqrt{\left(1 - \frac{1}{\tau\mu_{N-1}}\right)^2 - \frac{4c}{\tau\mu_{N-1}}}. \tag{14}$$

Moreover, $\tilde{y}(t^*)$ from (C6) obeys the Riccati differential equation (Appendix C):

$$\frac{d\tilde{y}}{dt^*} = (\tau\mu_{N-1} - 1)\tilde{y} - \tau\mu_{N-1}^2\tilde{y}^2 - c.$$

For the same initial $y(0) = \tilde{y}(0) = y_0$, the prevalence is bounded by the relatively simple expression (13)

$$\begin{cases} y(t^*) \geqslant \tilde{y}(t^*)|_{c=c_U} & \text{if } \Psi \leqslant c_U \\ y(t^*) \leqslant \tilde{y}(t^*)|_{c=c_L} & \text{if } \Psi \geqslant c_L \end{cases}.$$

Apart from the fact that $\tilde{y}(t^*)$ in (13) can lower and upper bound the SIS prevalence $y(t^*)$ as a function of the scaled time t^*, the formula (13) can be regarded as a time-dependent approximation for the SIS prevalence, when the constant c, appearing in Ξ, can be estimated or tuned in certain time intervals [e.g., when Ψ in (10) varies significantly with time]. The "tanh formula" $\tilde{y}(t^*)$ in (13) depends on the algebraic connectivity μ_{N-1} and the constant c, that both reflect properties of the underlying graph G. Appendix C demonstrates that $\lim_{t\to\infty}\tilde{y}(t) = \tilde{y}_\infty$, where

$$\tilde{y}_\infty = \frac{1}{2}\left(1 - \frac{1}{\tau\mu_{N-1}}\right) + \frac{1}{2}\sqrt{\left(1 - \frac{1}{\tau\mu_{N-1}}\right)^2 - \frac{4c}{\tau\mu_{N-1}}},$$

which only holds for certain values τ, depending on c. Indeed, the condition (C8) on c in Appendix C translates to

$$-1 \leqslant c \leqslant \frac{(\tau\mu_{N-1})}{4}\left(1 - \frac{1}{\tau\mu_{N-1}}\right)^2,$$

so that $c_L = -1$ and $c_U = \frac{(\tau\mu_{N-1})}{4}(1 - \frac{1}{\tau\mu_{N-1}})^2$. If $1 - \frac{1}{\tau\mu_{N-1}} < 0$ or $\tau < \frac{1}{\mu_{N-1}}$, then $c < 0$ because $\tilde{y}_\infty \geqslant 0$.

The lower bound $c_L \leqslant \Psi$ in the definition (10) leads to a bound for the effective infection rate

$$\tau_L \leqslant \frac{1}{\mu_{N-1}\left(\frac{E[R]}{N\mu_{N-1}} - \text{var}[S]\right)}. \tag{15}$$

The upper bound $\Psi \leqslant c_U$ implies that $y(t^*) \geqslant \tilde{y}(t^*)|_{c=c_U}$ and $\tilde{y}(t^*)|_{c=c_U}$ is real if the discriminant in (14) is positive:

$$\frac{1}{\mu_{N-1}\left(1 - 2\sqrt{\text{var}[S] - \frac{E[R]}{N\mu_{N-1}}}\right)} \leqslant \tau_U. \tag{16}$$

We observe that the bound (15) for τ_L is meaningless if $\text{var}[S] - \frac{E[R]}{N\mu_{N-1}} > 0$ or $\Psi > 0$, whereas, if $\text{var}[S] - \frac{E[R]}{N\mu_{N-1}} < 0$ or $\Psi < 0$, the bound (16) for τ_U is unrealistic.

A. Comparison of the tanh formula (13) with NIMFA

The N-intertwined mean-field approximation (NIMFA) has been proposed first in Ref. [16], but later in Ref. [5] a more elegant approach was found that we briefly repeat here. Introducing the covariance [15, p. 25], $\mathrm{Cov}[X_i,X_k] = E[X_iX_k] - E[X_i]E[X_k]$ in the governing equation (1) leads to

$$\frac{dE[X_i]}{dt} = -\delta E[X_i] + \beta(1-E[X_i])\sum_{k=1}^{N} a_{ki}E[X_k] - \beta\sum_{k=1}^{N} a_{ki}\mathrm{Cov}[X_i,X_k].$$

In terms of the infection probability of node i, $w_i(t) = \Pr[X_i(t)=1] = E[X_i(t)]$, we have

$$\frac{dw_i(t)}{dt} = -\delta w_i(t) + \beta[1-w_i(t)]\sum_{k=1}^{N} a_{ki}w_k(t) - \beta\sum_{k=1}^{N} a_{ki}\mathrm{Cov}[X_i(t),X_k(t)]. \quad (17)$$

The first part is recognized as the NIMFA equation [24] in the mean-field infection probability $v_i(t) = \Pr[X_i^{(1)}(t)=1]$ of node i, where $X_i^{(1)}$ is the first order (or NIMFA) mean-field approximation (MFA) of the state X_i of node i,

$$\frac{dv_i(t)}{dt} = -\delta v_i(t) + \beta[1-v_i(t)]\sum_{k=1}^{N} a_{ki}v_k(t). \quad (18)$$

For each node i, we may consider $\beta\tilde{R}_i$ in (17), where

$$\tilde{R}_i = \sum_{k=1}^{N} a_{ki}\mathrm{Cov}[X_i(t),X_k(t)] \quad (19)$$

as the MFA correction term, whose omission in (17) specifies the impact or accuracy (per node) of MFA. Clearly, if $\mathrm{Cov}[X_i(t),X_k(t)]=0$ for each nodal pair (i,k), then the NIMFA equations (18) are equal to the exact SIS equations (17). Moreover, as shown in Ref. [25], $\mathrm{Cov}[X_i(t),X_k(t)] \geqslant 0$ for a Markovian SIS and SIR process on any graph, so that $\beta\tilde{R}_i \geqslant 0$ and NIMFA always upper bounds the viral infection probability ($v_i \geqslant w_i$) and, thus, lower bounds the epidemic threshold $\tau_c \geqslant \tau_c^{(1)} = \frac{1}{\lambda_1}$, where λ_1 is the spectral radius of the adjacency matrix A of the graph [22].

We compare the NIMFA prevalence $y^{(1)}(t^*) = \frac{1}{N}\sum_{i=1}^{N} v_i(t^*)$, where $v_i(t)$ obeys the NIMFA equations (18), and our new approximation, called a "tanh formula" $\tilde{y}(t^*)$ in (13) for different values of c with simulations of the exact Markovian SIS process on the same graph.

Figure 1 illustrates the prevalence as a function of time $t^* = t$ (for $\delta = 1$) in one realization of an Erdős-Rényi random graph $G_p(N)$ with $N = 50$ nodes and link density $p = 0.4$, spectral radius $\lambda_1 = 20.85$ of the adjacency matrix, and algebraic connectivity $\mu_{N-1} = 10.11$. The normalized infection strength $x = \frac{\tau}{\tau_c^{(1)}} = \lambda_1\tau$ was chosen in Fig. 1 to operate below the epidemic threshold τ_c. Both NIMFA and the tanh formula in (13) for $\tilde{y}(t^*)$ with $c = 0$ are compared with simulations. Initially, only one node is infected, and 10^6 realizations of the Markovian SIS process are simulated to produce an accurate average fraction of infected nodes $y(t^*)$.

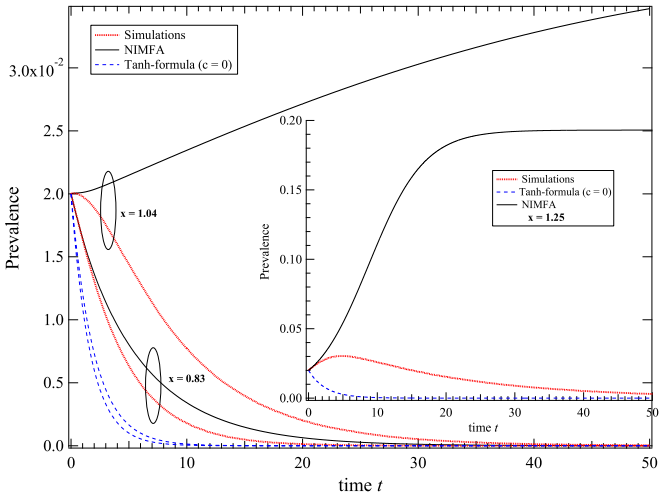

FIG. 1. The prevalence as a function of time $t^* = t$ (for $\delta = 1$) in one realization of an Erdős-Rényi random graph $G_p(N)$ with $N = 50$ nodes and link density $p = 0.4$, spectral radius $\lambda_1(A) = 20.85$, and algebraic connectivity $\mu_{N-1} = 10.11$. The normalized infection strength $x = \frac{\tau}{\tau_c^{(1)}} = \lambda_1(A)\tau$ was chosen to operate below the epidemic threshold τ_c. Both NIMFA and the tanh formula in (13) for $\tilde{y}(t^*)$ with $c = 0$ are compared with simulations. Initially, only one node is infected, and 10^6 realizations of the Markovian SIS process are simulated to produce an accurate average fraction of infected nodes $y(t^*)$.

the tanh formula in (13) for $\tilde{y}(t^*)$ with $c = 0$ are compared with simulations. Initially, only one randomly chosen node is infected, and 10^6 realizations of the Markovian SIS process are used to produce the average fraction of infected nodes $y(t^*)$.

Figure 2 shows the prevalence above the epidemic threshold τ_c, for precisely the same graph as in Fig. 1. The value of c in the tanh formula in (13) was chosen $c = -0.055$ (for $x = 2.08$) and $c = -0.06$ (for $x = 4.17$), respectively, in order to closely approach the simulated metastable prevalence.

Both Figs. 1 and 2 exhibit rather large deviations of both NIMFA and the tanh formula (13) during the transient phase (initial regime from $t = 0$ up to $t^* = t_m^*$, the start of the metastable regime). On the other hand, as is well known, the

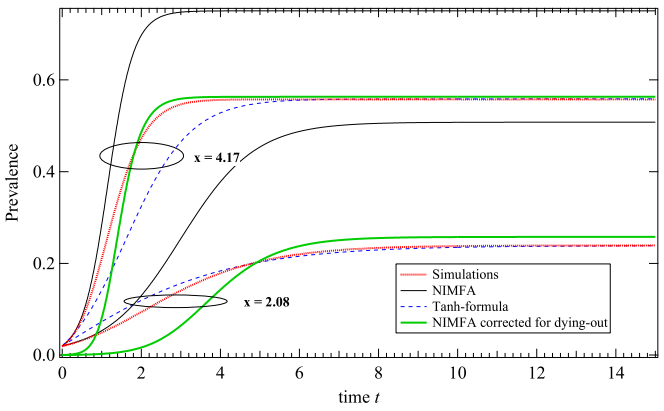

FIG. 2. The prevalence as a function of time $t^* = t$ (for $\delta = 1$) for precisely the same graph as in Fig. 1, with normalized infection strength $x = \frac{\tau}{\tau_c^{(1)}} = \lambda_1(A)\tau$ well above the epidemic threshold τ_c.

metastable regime, in which the prevalence is more or less constant, is much better approached by SIS approximations. So far, very few papers have investigated the time-dependent behavior of the prevalence on networks. Since initially one node is infected, in any graph there is a non-negligible chance that the outbreak sparked by this this initial node dies out, even for $\tau > \tau_c$. Moreover, the metastable regime is dependent on the initial condition (both the number of infected nodes as well as their location in the graph). We have corrected[1] NIMFA with the probability to die out in Fig. 2, so that the deviations between simulations and NIMFA are more reasonable.

IV. NEW UPPER BOUNDS FOR THE EPIDEMIC THRESHOLD FOR GENERAL GRAPHS

The Riccati approximation (13) increases with t^* as indicated by (C7) when the condition (C8) on c is satisfied. Due to the existence of the absorbing state (i.e., the overall-healthy state), $\lim_{t^*\to\infty} y(t^*) = 0$, implying that the bound $y(t^*) \geqslant \tilde{y}(t^*)|_{c=c_U}$ can apply only in the time interval $[0,t_m^*)$, where $t_m^* = \theta$ denotes the start of the metastable regime (for $\tau > \tau_c$) obeying $\frac{dy(t^*)}{dt^*}|_{t^*=t_m^*} = 0$, and where $y(t^*)$ is increasing. Thus, when replacing the upper bound $\Psi \leqslant c_U$, which can hold for any time point t^*, by $\max_{0\leqslant t^*\leqslant t_m^*} \Psi \leqslant c_U$, the bound (16) becomes

$$\frac{1}{\mu_{N-1}\left\{1 - 2\sqrt{\max_{0\leqslant t^*\leqslant t_m^*}\left(\mathrm{var}[S] - \frac{E[R]}{N\mu_{N-1}}\right)}\right\}} \leqslant \tau_U.$$

Since for $\tau > \tau_U$, it follows that $y(t^*) \geqslant \tilde{y}(t^*)|_{c=c_U} > 0$, meaning that there is in the graph a nonzero average fraction of infected nodes, and we find an upper bound for the epidemic threshold [defined as $\tau \leqslant \tau_c : y(t^*,\tau) = 0$ for sufficiently large t^*]

$$\tau_c \leqslant \frac{1}{\mu_{N-1}\left\{1 - 2\sqrt{\max_{0\leqslant t^*\leqslant t_m^*}\left(\mathrm{var}[S] - \frac{E[R]}{N\mu_{N-1}}\right)}\right\}}.$$

Since

$$\max_{0\leqslant t^*\leqslant t_m^*}\left(\mathrm{var}[S] - \frac{E[R]}{N\mu_{N-1}}\right) \leqslant \max_{0\leqslant t^*\leqslant t_m^*} \mathrm{var}[S]$$

and $\max_{0\leqslant t^*\leqslant t_m^*} \mathrm{var}[S] \leqslant \max_{t\geqslant 0} \mathrm{var}[S(t;\tau)]$, we arrive at the following:

Theorem 1. For any graph, the SIS epidemic threshold is upper bounded by

$$\tau_c \leqslant \frac{1}{\mu_{N-1}\{1 - 2\sqrt{\max_{t\geqslant 0} \mathrm{var}[S(t;\tau_c)]}\}}. \tag{20}$$

Unfortunately, the variance $\mathrm{var}[S(t;\tau)]$ of the number of infected nodes in the metastable state is difficult to determine, although Theorem 1 shows how essential $\mathrm{var}[S(t;\tau)]$ is. Equality in (20) holds for the complete graph K_N [see Ref. [15, p. 456] and (31) below] at an extremal time $t^* = \theta$. The upper bound $\tau_c \leqslant \frac{2}{\mu_{N-1}}$, proved by Ganesh *et al.* [26], indicates that, for K_N,

$$\mathrm{var}[S(\theta;\tau_c)] \leqslant \tfrac{1}{16},$$

which is substantially smaller than the maximum possible[2] variance of $\frac{1}{4}$. The sharpest possible upper bound occurs when $\max_{t^*} \mathrm{var}[S(t^*;\tau_c)] = 0$ in (20), leading to

$$\tau_c \leqslant \frac{1}{\mu_{N-1}}, \tag{21}$$

which is violated for the complete graph K_N, because there is only one graph, K_N, for which $\mu_{N-1} = N > \lambda_1 = N - 1$.

When combining the hypothetical upper bound (21) with the lower bound $\tau_c \geqslant \tau_c^{(1)} = \frac{1}{\lambda_1}$, then we would find that, for any graph except for the complete graph K_N, the SIS epidemic threshold is bounded by

$$\frac{1}{\lambda_1} \leqslant \tau_c \leqslant \frac{1}{\mu_{N-1}}. \tag{22}$$

For dense graphs, where both the spectral radius λ_1 and the algebraic connectivity μ_{N-1} are large, the difference $\frac{1}{\mu_{N-1}} - \frac{1}{\lambda_1}$, can be small, whereas roughly the opposite holds for sparse graphs. For example, for a cycle C, the lower and upper bound in (22) lie far apart: the spectral radius is $\lambda_1 = 2$ and the algebraic connectivity [22, p. 123] equals $(\mu_C)_{N-1} = 4\sin^2\frac{\pi}{N}$, so that $\frac{1}{(\mu_C)_{N-1}} = O(N^2)$ for large N, whose scaling with N is substantially larger than the epidemic threshold deduced in Refs. [2,27,28]. While the upper bound (22) is correct in many sparse graphs, we now demonstrate that it cannot hold in general (even excluding K_N). If the upper bound (22) is true, then there would exist finite graphs for which the spectral radius λ_1 equals the algebraic connectivity μ_{N-1}, and, for those graphs, (22) would return the exact SIS epidemic threshold, for any size N. For those graphs, which we call "$\lambda_1 = \mu_{N-1}$ graphs," the NIMFA epidemic threshold $\tau_c^{(1)}$ is also exact. For any graph excluding K_N, it holds that $\mu_{N-1} \leqslant d_{\min}$ (see, e.g., Ref. [22, p. 82]), while $d_{\min} \leqslant d_{av} \leqslant \lambda_1$. For any nonregular graph, the minimum degree is strict smaller than the average degree, $d_{\min} < d_{av}$, so that $\mu_{N-1} \leqslant d_{\min} < \lambda_1$. Hence, there

[1] When only node i is infected in the graph G at a certain time, the continuous-time Markovian SIS epidemic process has two possibilities: die-out with curing rate δ or grow to two infected nodes, when node i can infect one of its d_i neighbors. Hence, the die-out situation at node i is described by a two-state continuous-time Markov process with transition rate from one to zero infected nodes equal to $q_{10} = \delta$ and from one to two infected nodes equal to $q_{12} = \beta d_i$, from which we find [15, pp. 220–221] that, given only node i is infected, the probability that the epidemic dies out is $\Pr[\text{die-out}] = \frac{q_{01}}{q_{01}+q_{12}} = \frac{\delta}{\delta+\beta d_i} = \frac{1}{1+\tau d_i}$. When initially more than one node is infected, the die-out probability rapidly becomes negligibly small (for $\tau > \tau_c$). As an approximation, we have replaced here in Fig. 2 the degree of node i by its average $E[D] = p(N-1) \approx \lambda_1$ so that $\Pr[\text{die-out}] \approx \frac{1}{1+x}$, where $x = \lambda_1\tau$. The probability to reach the metastable state in NIMFA, corrected for die-out, is $y^{(1)}(\tau)\Pr[\text{no die-out}] \approx y^{(1)}(\tau)(1 - \frac{1}{1+x})$. Finally, in a regular graph with degree $r = \lambda_1$, the NIMFA prevalence is $y^{(1)}(\tau) = 1 - \frac{1}{x}$ and the corrected prevalence (to prevent die out when initially one node is infected) is about $(1 - \frac{1}{x})^2$.

[2] Since the random variable $S \in [0,1]$ in (2), we find, for all $p \geqslant 1$, that $S^p \leqslant S$. The variance of S, $\mathrm{var}[S] = E[S^2] - (E[S])^2$, is upper bounded, using $E[S^p] \leqslant E[S]$ for $p = 2$, by $\mathrm{var}[S] \leqslant y - y^2$, where $y = E[S]$, as defined in (3). The maximum of the function $y - y^2$ equals $\frac{1}{4}$ and occurs at $y = \frac{1}{2}$, and hence, $\mathrm{var}[S] \leqslant \frac{1}{4}$.

do not exist "$\lambda_1 = \mu_{N-1}$ nonregular graphs." On the other hand, regular graphs with degree r possess a spectral gap $\lambda_1 - \lambda_2 = \mu_{N-1}$ and $\lambda_1 = r$ is an integer. Then "$\lambda_1 = \mu_{N-1}$ regular graphs" must possess a zero second eigenvalue of the adjacency matrix, $\lambda_2 = 0$, which is satisfied only for complete multipartite graphs. In general but excluding K_N, the second largest adjacency eigenvalue is positive, $\lambda_2 > 0$, if and only if the graph is *not* complete multipartite [29, Sec. 3.4]). For regular complete multipartite graphs, all of its parts must to be equal. In conclusion, the only "$\lambda_1 = \mu_{N-1}$ regular graphs" are regular complete multipartite graphs of the form $K_{\{n_1,n_2,\ldots,n_m\}}$, where each part $n_1 = n_2 = \cdots = n_m = n$ and $nm = N$. An example is the regular complete bipartite graph $K_{n,n}$ with $2n = N$, whose NIMFA prevalence y_∞ is analytically known [16] (and that, as a curiosity, represents the exact SIS Markov graph as shown in Ref. [30] with $n = 2^{N-1}$). Physically, as mentioned earlier, the SIS epidemic threshold of the phase transition is only "zero-one sharp" when $N \to \infty$ and any finite graph will exhibit a small (but nonzero) τ region that characterizes the transition in the prevalence. In summary, the hypothetical upper bound (21) is not correct for finite "$\lambda_1 = \mu_{N-1}$ graphs," and likely also not in other dense graphs.

We provide another upper bound for the SIS epidemic threshold, by invoking the Motzkin-Straus Theorem [31] stating that

$$\left(1 - \frac{1}{\omega}\right) = \max_{x \in \mathcal{S}} x^T A x, \tag{23}$$

where the simplex $\mathcal{S}$ contains all vectors x that lie in the hyperplane $u^T x = 1$ and possesses non-negative components and where ω is the clique number in a graph G, defined as the size of the largest clique in G. Evidently, $\omega_{K_N} = N$ and in any other graph than K_N, the clique size $\omega \leqslant d_{\max}$.

Theorem 2. The SIS epidemic threshold is upper bounded by

$$\tau_c \leqslant \frac{1}{d_{\min} - N\left(1 - \frac{1}{\omega}\right)\sqrt{\max_{t \geqslant 0} \mathrm{var}[S(t, \tau_c)]}} \tag{24}$$

in any graph, for which the right-hand side in (24) is positive.

Theorem 2 is proved in Appendix D 2. The upper bound in (24) is minimized when $\mathrm{var}[S] = 0$,

$$\tau_c \leqslant \frac{1}{d_{\min}}. \tag{25}$$

For the cycle (or ring) with $\omega = d_{\min} = 2$, the upper bound (25) does not apply (see Sec. V C), implying that there must exist a value of $\tau > \frac{1}{d_{\min}} = \frac{1}{2}$ for which $\mathrm{var}[S_C] > \frac{1}{N^2}\left(1 - \frac{1}{\tau/2}\right)^2$ for sufficiently large N. The maximum possible bound $\mathrm{var}[S] = \frac{1}{4}$ in (24) yields

$$\tau_c \leqslant \frac{1}{d_{\min} - \frac{N}{2}\left(1 - \frac{1}{\omega}\right)} \tag{26}$$

for dense graphs where $d_{\min} > \frac{N}{2}\left(1 - \frac{1}{\omega}\right)$. In particular, for the complete graph K_N where $\omega = N$, (26) becomes $(\tau_c)_{K_N} \leqslant \frac{2}{N-1}$, which is not bad at all!

Since $\mu_{N-1} \leqslant d_{\min}$ (excluding K_N), the upper bound (25) is sharper than (21). However, in contrast to (21), the more confining condition on $\mathrm{var}[S] \leqslant \left(\frac{d_{\min}}{2N(1-\frac{1}{\omega})}\right)^2 \left(1 - \frac{1}{d_{\min}\tau}\right)^2$ in the proof of Theorem 2 must be satisfied.

In conclusion, Theorems 1 and 2 return valuable bounds on the SIS epidemic threshold in *dense graphs*. Unfortunately, for *sparse graphs*, most real-world networks are sparse, the general theory is hardly useful, illustrating the need for sophisticated methods (as developed, e.g., by Liggett [2,32] for infinite d lattices).

V. REGULAR GRAPHS

For regular graphs with degree r, the maximum average fraction of infected nodes equals[3] [15, p. 456]

$$y_{\infty;\,\mathrm{regular}}(\tau) = \frac{1}{N} \frac{E\left[w_\infty^T A w_\infty\right]}{r - \frac{1}{\tau}}, \tag{27}$$

where w_∞ corresponds to the extremal point θ at which $\frac{dy(t^*;\tau)}{dt^*}|_{t^*=\theta} = 0$, explained in Sec. II. For sufficiently large N and $\tau > \tau_c$, (27) accurately approximates the average fraction of infected nodes in the metastable state of the SIS process. Equation (27) shows that $y_{\infty;\,\mathrm{regular}}(\tau) \geqslant 0$ only if $\tau \geqslant \frac{1}{r}$, which corresponds to the NIMFA lower bound for the epidemic threshold. The NIMFA prevalence, denoted by a superscript $^{(1)}$, equals [15, p. 465]

$$y^{(1)}_{\infty;\,\mathrm{regular}}(\tau) = 1 - \frac{1}{r\tau}.$$

A. The cycle (or ring)

The SIS epidemic threshold $(\tau_c)_C$ in the ring is still unknown [33]. As reported by Neal [34], the best bounds so far for an infinitely large ring, are $1.539 \leqslant (\tau_c)_C \leqslant 1.942$, due to Liggett [2], where the upper bound 1.942 is the largest zero [2] of the polynomial $4x^3 - 7x^2 - 2x + 1$ and the lower bound 1.539 is much better than the NIMFA lower bound $(\tau_c^{(1)})_C = \frac{1}{2}$. Via a new method [28] based on the survival time, the epidemic threshold of the ring or cycle C (with degree $r = d_{\min} = 2$) is shown to be $(\tau_c)_C = \frac{f_C(N)}{d_{\min}}$, where $f_C(N)$ is a slowly increasing function of N. In particular, $f_C(500) \approx 2.8$ and $f_C(1000) \approx 3$. By extensive simulations, Ricardo and de Mendonca [27] found the estimate $\lim_{N\to\infty} \frac{f_C(N)}{2} \approx 1.649$, which is remarkably close to the Riemann-zeta function $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ evaluated at $s = 2$, namely, $\zeta(2) = \frac{\pi^2}{6} \approx 1.645$.

Theorem 3. The SIS epidemic threshold in the cycle C (or ring) is

$$(\tau_c)_C = \frac{1}{2(1 - 2\sqrt{\xi})}, \tag{28}$$

where $\xi = \mathrm{Cov}[X_{1\infty}, X_{2\infty}]$ is the SIS covariance between two neighboring nodes in the metastable state.

Proof. The adjacency matrix for the ring or cycle C is $A_C = E + E^{-1}$, where E is the elementary circulant

[3] Since $Q = rI - A$, (4) reduces to $y_\infty = \frac{\tau}{N} E[w_\infty^T r w_\infty - w_\infty^T A w_\infty]$. With $E[w_\infty^T r w_\infty] = rNE[S_\infty] = rNy_\infty$, we arrive at (27).

matrix [22, p. 116], so that

$$w^T A w = \sum_{i=1}^{N}\sum_{j=1}^{N} X_i (A_C)_{ij} X_j$$
$$= \sum_{i=1}^{N} X_i X_{(i+1) \bmod N} + \sum_{i=1}^{N} X_i X_{(i-1) \bmod N},$$

and $0 \bmod N$ is here equal to N. Further, in the metastable state, we have that, due to symmetry,

$$E[X_i X_{(i+1) \bmod N}] = E[X_i X_{(i-1) \bmod N}] = E[X_{1\infty} X_{2\infty}],$$

for each node i, so that $E[w_\infty^T A w_\infty] = 2N E[X_{1\infty} X_{2\infty}]$. Rewritten with $E[X_{1\infty} X_{2\infty}] = E[X_{1\infty}]E[X_{2\infty}] + \xi$, where $\xi = \mathrm{Cov}[X_{1\infty}, X_{2\infty}]$, and with $y_{\infty;C} = E[X_{1\infty}] = E[X_{2\infty}]$, as

$$E\left[w_\infty^T A w_\infty\right] = 2N\left[y_{\infty;C}^2(\tau) + \xi\right]$$

and substituted into (27) yields

$$y_{\infty;C}(\tau) = \frac{2\left(y_{\infty;C}^2(\tau) + \xi\right)}{2 - \frac{1}{\tau}}$$

from which

$$y_{\infty;C}^2(\tau) - \left(1 - \frac{1}{2\tau}\right) y_{\infty;C}(\tau) + \xi = 0.$$

Solving the quadratic equation, taking into account that $y_{\infty;C}(0) = 0$ and $\lim_{\tau\to\infty} y_{\infty;C}(\tau) = 1$, yields

$$y_{\infty;C}(\tau) = \frac{1}{2}\left(1 - \frac{1}{2\tau}\right) + \sqrt{\left[\frac{1}{2}\left(1 - \frac{1}{2\tau}\right)\right]^2 - \xi}. \quad (29)$$

Similarly as for the complete graph K_N, derived in Ref. [15, p. 457], a positive discriminant is required for a realistic $y_{\infty;C}(\tau) \geqslant 0$, leading to (28). ■

It remains to find an analytic expression for $\xi = \mathrm{Cov}[X_{1\infty}, X_{2\infty}]$ in terms of N. However, using the estimate $(\tau_c)_{\text{ring}} = 1.64896$ of Ref. [27] in (28) leads to $\xi = 0.121375$, which is about half of the maximum[4] possible $1/4$.

B. Application of the Motzkin-Straus Theorem

Invoking (B16) deduced from the Motzkin-Straus Theorem [31] in Appendix B into (27) yields

$$y_{\infty;\,\text{regular}}(\tau) \leqslant N \frac{1 - \frac{1}{\omega}}{r - \frac{1}{\tau}} E\left[S_\infty^2\right],$$

which is, with $E[S_\infty^2] = y_{\infty;\,\text{regular}}^2(\tau) + \mathrm{var}[S_\infty]$, rewritten as

$$y_{\infty;\,\text{regular}}^2(\tau) - \frac{\left(r - \frac{1}{\tau}\right)}{N\left(1 - \frac{1}{\omega}\right)} y_{\infty;\,\text{regular}}(\tau) + \mathrm{var}[S_\infty] \geqslant 0.$$

Provided the discriminant is non-negative, equivalent to

$$\mathrm{var}[S_\infty] \leqslant \left[\frac{\left(r - \frac{1}{\tau}\right)}{2N\left(1 - \frac{1}{\omega}\right)}\right]^2$$

and which agrees with Theorem 2, the solution of the quadratic inequality (using similar arguments as for the ring above) yields

$$y_{\infty;\,\text{regular}}(\tau) \geqslant \frac{r}{2N\left(1 - \frac{1}{\omega}\right)}\left(1 - \frac{1}{r\tau}\right) + \sqrt{\left[\frac{r}{2N\left(1 - \frac{1}{\omega}\right)}\left(1 - \frac{1}{r\tau}\right)\right]^2 - \mathrm{var}[S_\infty]}. \quad (30)$$

The lower bound (30) for the prevalence is sharpened in two extreme cases: (a) the complete graph K_N [15, p. 456]

$$y_{\infty;K_N}(\tau) = \frac{1}{2}\left(1 - \frac{1}{N\tau}\right) + \sqrt{\left[\frac{1}{2}\left(1 - \frac{1}{N\tau}\right)\right]^2 - \mathrm{var}[S_{\infty;K_N}]} \quad (31)$$

and (b) the ring in (29).

For $\tau \to \infty$, where $y_{\infty;\,\text{regular}}(\tau) \to 1$ and $\mathrm{Var}[S_\infty] \to 0$, the inequality (30) leads to Wilf's bound [35] for the clique number, $\omega \geqslant \frac{1}{1 - \frac{\lambda_1}{N}}$ for $\lambda_1 = r$, in which equality is reached for the complete graph K_N.

C. Brief overview on upper bounds for τ_c

To the best of our knowledge, very few general (and nontrivial) upper bounds exist for the SIS epidemic threshold τ_c. Boguñá *et al.* [36] have proposed a physically relevant procedure to find a tight upper bound for τ_c; however, their upper bound is not mathematically rigorous. Ganesh *et al.* [26] demonstrated that $\tau_c \leqslant \frac{1}{\eta}$, where the isoperimetric constant η is related [22, p. 95] to the algebraic connectivity μ_{N-1} by $\eta \geqslant \frac{\mu_{N-1}}{2}$, so that $\tau_c \leqslant \frac{1}{\eta} \leqslant \frac{2}{\mu_{N-1}}$. Hence, their upper bound is better than $\frac{2}{\mu_{N-1}}$.

When rewriting (8) as

$$\tau^{-1} = \frac{E[w(\theta;\tau)^T Q w(\theta;\tau)]}{N y(\theta;\tau)},$$

we can define the epidemic threshold τ_c as that value of τ for which the order parameter $y(\theta;\tau)$ approaches zero from above:

$$\tau_c^{-1} = \lim_{y(\theta;\tau)\downarrow 0} \frac{E[w(\theta;\tau)^T Q w(\theta;\tau)]}{N y(\theta;\tau)}.$$

An upper bound for the SIS epidemic threshold, proven in Ref. [4], is the following:

Theorem 4. Let

$$\varepsilon_G = \lim_{y_\infty \downarrow 0} \max_{(k,l)\in\mathcal{L}} \Pr[X_k = 1 | X_l = 1],$$

where the involved quantities such as the prevalance y_∞ and the node k's infectious state X_k are computed in the metastable

[4] As shown in [15, p. 26], any pair of random variables X and Y obeys the inequality $\mathrm{Cov}[X,Y] \leqslant \sqrt{\mathrm{var}[X]\mathrm{var}[Y]}$. Since $X_{1\infty}$ and $X_{2\infty}$ possess the same distribution, we arrive at $\mathrm{Var}[X] \leqslant \frac{1}{4}$ for a Bernoulli random variable X.

state. Then the SIS epidemic threshold τ_c in graph G is upper bounded by

$$\tau_c \leqslant \frac{1}{d_{\min}(1-\varepsilon_G)}. \tag{32}$$

Theorem 4 emphasizes, via ε_G, the role of the joint probability of infection at end nodes of a same link. Earlier in Refs. [15, p. 458] and [4], unfortunately, an error has been made, which we would like to correct here. The argument that "the conditional probability ε_G in Theorem 4 can be upper bounded by $\varepsilon_G \leqslant \varepsilon_{K_N}$, because just at the onset of infection ($y \downarrow 0$), the maximum conditional infection probability ε_G on a link (k,l) in the graph G is largest in the complete graph K_N" is false (it should be "smallest"), as well as the ensuing upper bound based on K_N, for any graph,

$$\tau_c \leqslant \frac{1}{d_{\min}}\left[1 + O\left(\frac{1}{\sqrt{N}}\right)\right]. \tag{33}$$

In other words, in some regular graphs such as the cycle (see Theorem 3), the epidemic threshold τ_c can be factors larger than $\frac{1}{d_{\min}}$, contradicting (33).

VI. SUMMARY

A new approximation $\tilde{y}(t^*)$ in (13) for the time-varying SIS prevalence has been proposed, which seems accurate, provided the value of c in Ξ in (14) can be estimated closely. The analytic nature of the new approximation $\tilde{y}(t^*)$ in (13) allows very fast computations of the prevalence $y(t^*)$ at each time t^* and may serve as an approximate tool to model time-dependent aspects, such as the average time to reach the metastable state, given some initial infected fraction $y(0)$. A deeper investigation of the tanh formula in (13), in comparison with NIMFA, and of the function Ψ in (10) for different types of graphs stands on the agenda of future work.

After a study of the time dependence of the SIS prevalence, two general upper bounds for the SIS epidemic threshold [(21) and (24)] and specific ones for regular graphs are added that, unfortunately, still contain an unknown variance, which cannot be ignored in general (see also Lemma 1 and Ref. [25]). The differential equation of var[S], derived in Ref. [37],

$$\frac{d}{dt^*}\text{var}[S] = -2\text{var}[S] + \frac{2\tau}{N}E[(S-E[S])w^T Q w] + \frac{1}{N}\left(y + \frac{\tau}{N}E[w^T Q w]\right), \tag{34}$$

may lead to further progress, although third order moments in S appear.

ACKNOWLEDGMENTS

We are grateful to Dragan Stevanović for pointing that the only "$\lambda_1 = \mu_{N-1}$ regular graphs" are regular complete multipartite graphs, to Stojan Trajanovski, Jil Meier for spotting errors in an earlier version, and to Qiang Liu for providing data for Figs. 1 and 2.

APPENDIX A: PROOF OF THE BASIC DIFFERENTIAL EQUATION (4)

Summing the SIS governing equation (1) over all nodes i and omitting the time-dependence in $X_i(t)$ to shorten the equations yields

$$\frac{dE\left[\sum_{i=1}^N X_i\right]}{dt} = E\left[-\delta\sum_{i=1}^N X_i + \beta\sum_{k=1}^N\sum_{i=1}^N a_{ki}X_k - \beta\sum_{k=1}^N\sum_{i=1}^N a_{ki}X_iX_k\right],$$

and, after letting $t^* = t\delta$ and $\tau = \frac{\beta}{\delta}$ and using that the degree $d_k = \sum_{i=1}^N a_{ki}$, and written in terms of the prevalence (3), we obtain

$$\frac{dy(t^*;\tau)}{dt^*} = -y(t^*;\tau) + E\left[\frac{\tau}{N}\sum_{k=1}^N d_kX_k - \frac{\tau}{N}\sum_{k=1}^N\sum_{i=1}^N a_{ki}X_iX_k\right].$$

If we define the vector $w(t) = (X_1, X_2, \ldots, X_N)$ of random variables and the degree vector D with ith vector component the degree d_i of node i, then

$$\frac{dy(t^*;\tau)}{dt^*} = -y(t^*;\tau) + \frac{\tau}{N}E[D^T w(t^*) - w^T(t^*)Aw(t^*)]. \tag{A1}$$

We write the degree vector as $D = \Delta u$, where $\Delta = \text{diag}(d_1, d_2, \ldots, d_N)$ so that

$$\begin{aligned} D^T w - w^T Aw &= u^T\Delta w + w^T\Delta w - w^T\Delta w - w^T Aw \\ &= (u-w)^T\Delta w + w^T(\Delta - A)w. \end{aligned}$$

Now, $u - w$ is the (random) vector with noninfected nodes and

$$(u-w)^T\Delta w = \sum_{j=1}^N (1-X_j)d_jX_j = \sum_{j=1}^N \left(X_j - X_j^2\right)d_j = 0$$

because $X_j = X_j^2$ as $X_j \in \{0,1\}$. Introducing the Laplacian $Q = \Delta - A$ of the graph G, we arrive at the differential equation (4) for the average fraction of infected nodes y expressed as a quadratic form of the Laplacian Q. ■

APPENDIX B: THE QUADRATIC FORM $z^T Qz$

An analysis of the Laplacian quadratic form in an interconnected (or multilayer) network is presented in Ref. [23].

1. z is a real vector

Since the eigenvectors of Q constitute an orthogonal basis, we can write any $N \times 1$ real vector z as a linear combination of eigenvectors $x_1, x_2, \ldots, x_N$ of Q,

$$z = \sum_{k=1}^N \alpha_k x_k,$$

where $\alpha_k = z^T x_k$. Then

$$z^T Qz = \sum_{k=1}^N\sum_{m=1}^N \alpha_k\alpha_m x_k^T Qx_m = \sum_{k=1}^{N-1}\alpha_k^2\mu_k \tag{B1}$$

because the Laplacian eigenvalue $\mu_N = 0$ and

$$z^T z = \sum_{k=1}^{N} \alpha_k^2 = \frac{1}{N}(u^T z)^2 + \sum_{k=1}^{N-1} \alpha_k^2. \tag{B2}$$

Since Q is positive semidefinite (all eigenvalues $\mu_k \geqslant 0$), we find from (B1) the set of lower bounds

$$z^T Q z \geqslant \sum_{k \in K} \alpha_k^2 \mu_k,$$

where K is a set with at most $N-1$ different integers $k \in [1, N-1]$, and also,

$$\mu_{N-1} \sum_{k=1}^{N-1} \alpha_k^2 \leqslant z^T Q z \leqslant \mu_1 \sum_{k=1}^{N-1} \alpha_k^2,$$

which leads, with (B2), to

$$\mu_{N-1} \leqslant \frac{z^T Q z}{z^T z - \frac{1}{N}(u^T z)^2} \leqslant \mu_1 \tag{B3}$$

with equality in the upper bound for $z = x_1$ and in the lower bound for $z = x_{N-1}$. The lower bound is essentially the Rayleigh inequality [22, p. 73] for the algebraic connectivity μ_{N-1}.

2. w is a binary vector

The vector w is a so-called binary vector, because each component $w_k = X_k$ is either zero or one. For such vectors, we observe with (2) that

$$w^T w = \sum_{k=1}^{N} X_k^2 = \sum_{k=1}^{N} X_k = u^T w = NS.$$

Let us consider the eigenvector decomposition

$$w = \sum_{k=1}^{N} \zeta_k x_k, \tag{B4}$$

where $\zeta_k = w^T x_k$ and, explicitly,

$$\zeta_k = \sum_{j=1}^{N} X_j (x_k)_j = \sum_{n \in [1,N]: X_n = 1} (x_k)_n \tag{B5}$$

$$= - \sum_{m \in [1,N]: X_m = 0} (x_k)_m = - \sum_{j=1}^{N} (1 - X_j)(x_k)_j,$$

where the latter follows from $u^T x_k = 0$ for $1 \leqslant k < N$. When $X_j = 1$ or $X_j = 0$ for all j, then $\zeta_k = 0$ for $1 \leqslant k < N$. The particular case for $k = N$ with eigenvector vector $x_N = \frac{u}{\sqrt{N}}$ equals

$$\zeta_N = w^T x_N = \frac{1}{\sqrt{N}} w^T u = \sqrt{N} S$$

and

$$w = S\, u + \sum_{k=1}^{N-1} \zeta_k x_k, \tag{B6}$$

where the average over all nodes of $\sum_{k=1}^{N-1} \zeta_k x_k = 0$ (due to orthogonality $u^T x_k = 0$ for $1 \leqslant k < N$). After taking the expectation of both sides, we find for the j-component (corresponding to node j) that

$$E[X_j] = y + \sum_{k=1}^{N-1} E[\zeta_k](x_k)_j,$$

which expresses the nodal probability of infection $E[X_j] = \Pr[X_j = 1]$ in terms of the average fraction y of infected nodes (at zero frequency $\mu_N = 0$) and a correction due to the higher eigenfrequencies ($\mu_k > 0$).

Applying (B2) yields[5]

$$\sum_{k=1}^{N-1} \zeta_k^2 = u^T w - \frac{1}{N}(u^T w)^2 = N(S - S^2), \tag{B7}$$

which demonstrates that $\sum_{k=1}^{N-1} \zeta_k^2$ tends to zero for $S \to 0$ as well as for $S \to 1$. Since $S \in [0,1]$ so that $S - S^2 \leqslant \frac{1}{4}$, we find that $(\zeta_k^2)_{\text{av}} = \frac{1}{N} \sum_{k=1}^{N-1} \zeta_k^2 \leqslant \frac{1}{4}$, indicating that the arithmetic average squared coefficient $(\zeta_k^2)_{\text{av}}$ is rather small, especially in comparison with $\zeta_N^2 = NS^2$. Only a little can be said[6] about the expectation $E[\zeta_k^2]$.

Next, the quadratic form (B1) becomes

$$w^T Q w = \sum_{k=1}^{N-1} \zeta_k^2 \mu_k \tag{B8}$$

and shows, since all Laplacian eigenvalues $\mu_k > 0$ for $1 \leqslant k \leqslant N-1$ in a connected graph, that

$$w^T Q w \geqslant 0, \tag{B9}$$

where equality is possible only if all $\zeta_k^2 = 0$. In view of (B7), only if all $X_i = 0$ (corresponding to the overall healthy or absorbing state and $w = 0$) or all $X_i = 1$ (all nodes are infected and $w = u$), then $w^T Q w = 0$, but $w^T Q w > 0$ in any other situation. Invoking (B3) yields the lower bound

$$w^T Q w \geqslant \mu_{N-1} N(S - S^2), \tag{B10}$$

while the upper bound

$$w^T Q w \leqslant \mu_1 N(S - S^2) \tag{B11}$$

[5]Since the complement Bernoulli vector $w^c = u - w$ is orthogonal to w, (B7) also follows after executing the scalar product $w^T w^c = 0$ using (B6).

[6]We can show that $\sum_{k=1}^{N} (E[\zeta_k])^2 = \sum_{m=1}^{N} (E[X_m])^2$.

seems rather weak,[7] because w has no negative components so that w is more aligned to $x_N = u$ than to any other eigenvector of Q. Introducing from (B7)

$$\zeta_{N-1}^2 = N(S - S^2) - \sum_{k=1}^{N-2} \zeta_k^2$$

yields

$$w^T Q w = \mu_{N-1} N(S - S^2) + R,$$

where the correction R in (11) is

$$R = \sum_{k=1}^{N-2} (\mu_k - \mu_{N-1}) \zeta_k^2.$$

Alternatively, after Abel summation [22, p. 56], we obtain

$$w^T Q w = \sum_{k=1}^{N-2} \left(\sum_{j=1}^{k} \zeta_j^2 \right) (\mu_k - \mu_{k+1}) + \mu_{N-1} \sum_{j=1}^{N-1} \zeta_j^2.$$

Using (B7) shows that

$$w^T Q w = \mu_{N-1} N(S - S^2) + R, \tag{B12}$$

where another expression of the correction R is

$$R = \sum_{k=1}^{N-2} \left[\sum_{j=1}^{k} \zeta_j^2(t^*) \right] (\mu_k - \mu_{k+1}). \tag{B13}$$

Since the eigenvalue gaps $\mu_k - \mu_{k+1} \geqslant 0$ and $\sum_{j=1}^{k} \zeta_j^2 \geqslant 0$, each term in the sum in (B13) as well as in (11) is non-negative, demonstrating that $R \geqslant 0$. Only for the complete graph K_N are the nonzero Laplacian eigenvalues all equal to $\mu_k = N$ for $1 \leqslant k < N$, so that $R_{K_N} = 0$. Moreover, the complete graph K_N is the only graph that has two distinct eigenvalues [22, p. 45]. For all other graphs, $R > 0$ provided $0 < S < 1$, because $\zeta_k = 0$ for $1 \leqslant k < N$ if $S = 1$ or $S = 0$. Applying the upper bound (B11) and $w^T Q w \leqslant L$ to (B12) shows that

$$R \leqslant \min\left(\{\mu_1 - \mu_{N-1}\} N(S - S^2), L - \mu_{N-1} N(S - S^2)\right).$$

Sharper upper bounds of R for an arbitrary graph are currently not available.

We rewrite (4) with $y = E[S]$ and $S = \frac{1}{N} w^T w$ as

$$\frac{dy(t^*;\tau)}{dt^*} = \frac{1}{N} E[w(t^*;\tau)^T (\tau Q - I) w(t^*;\tau)], \tag{B14}$$

which is the expectation of a quadratic form of the matrix $\tau Q - I$. Introducing the spectral decomposition of the Laplacian $Q = \sum_{k=1}^{N} \mu_k x_k x_k^T$ yields

$$\frac{dy}{dt^*} = \frac{1}{N} \sum_{k=1}^{N} (\tau \mu_k - 1) E\left[\left(x_k^T w \right)^2 \right]$$

$$= \frac{1}{N} \sum_{k=1}^{N} (\tau \mu_k - 1) E\left[\zeta_k^2 \right].$$

With $\zeta_N = \sqrt{N} S$, we have

$$\frac{dy}{dt^*} = \frac{1}{N} \sum_{k=1}^{N-1} (\tau \mu_k - 1) E\left[\zeta_k^2 \right] - E[S^2]. \tag{B15}$$

3. A lower bound from the Motzkin-Straus Theorem

The condition $u^T x = 1$ in the Motzkin-Straus Theorem requires a scaling of $w = \zeta x$, so that $NS = w^T u = \zeta x^T u = \zeta$ and with (23), we have

$$w^T A w = N^2 S^2 x^T A x \leqslant N^2 S^2 \left(1 - \frac{1}{\omega} \right), \tag{B16}$$

which leads to the lower bound:

$$w^T Q w = w^T \Delta w - w^T A w \geqslant \sum_{i=1}^{N} d_i X_i - N^2 S^2 \left(1 - \frac{1}{\omega} \right). \tag{B17}$$

APPENDIX C: A RICCATI DIFFERENTIAL EQUATION

Consider the Riccati differential equation

$$\frac{dy}{dt} = ay - by^2 - g(t), \tag{C1}$$

where we assume that $g(t)$ is an arbitrary function, not depending on y. This differential equation has been solved in [38], when a and b are time-dependent functions, but where $g(t) = 0$, to model SIS epidemics on the complete graph with time-varying rate functions. In their seminal paper [20], Kermack and McKendrick exactly solve the differential equations for SIR epidemics in a homogeneous population (i.e., complete graph) with constant rates β and δ

$$\frac{dx}{dt} = -\beta x y \quad \frac{dy}{dt} = \beta x y - \delta y \quad \frac{dz}{dt} = \delta y,$$

where x, y, z denotes the number of susceptible, infected and removed items in the population of size $N = x + y + z$. A key observation in Ref. [20] is that

$$\frac{dx}{dz} = -\tau x$$

whose solution is $\log \frac{x(t)}{x_0} = -\tau z(t)$ [since $z(0) = 0$]. Writing $y = N - x - z$ in the last SIR differential equation and introducing $x = x_0 e^{-\tau z}$ yields

$$\frac{dz}{dt} = \delta(N - x_0 e^{-\tau z} - z) \tag{C2}$$

and integrated with $t^* = \delta t$

$$t^* = \int_{z_0}^{z} \frac{du}{N - x_0 e^{-\tau u} - u}.$$

Since the integral is not analytically known, Kermack and McKendrick approximate $e^{-\tau z} = 1 - \tau z + \frac{1}{2} \tau^2 z^2 + O(z^3)$ in (C2) to obtain

$$\frac{dz}{dt^*} = N - x_0 + (x_0 \tau - 1) z - \frac{x_0 \tau^2}{2} z^2,$$

which is of the form (C1) with constant $g(t) = N - x_0$, whose solution (C6) appears in Ref. [20] and is reviewed in Ref. [8, Sec. 2.3].

[7] We can prove another upper bound $w^T Q w \leqslant L$.

A Ricatti differential equation can be reduced to a second order linear differential equation. For, let $y = \alpha v$ and (C1) becomes

$$\frac{dv}{dt} = av - b\alpha v^2 - \frac{g(t)}{\alpha},$$

and we choose $b\alpha = 1$, so that $\alpha = \frac{1}{b}$ and

$$\frac{dv}{dt} + v^2 = av - bg(t).$$

Let $v = \frac{u'}{u}$, then $v' = \frac{u''}{u} - \frac{(u')^2}{u^2}$ so that $v' + v^2 = \frac{u''}{u}$, and we obtain the linear, second order differential equation

$$u'' - au' + bg(t)u = 0 \tag{C3}$$

with

$$y = \frac{1}{b}\frac{u'}{u}. \tag{C4}$$

It is questionable whether a general solution of the differential equation (C3) for an arbitrary function $g(t)$ can be found in analytic form. However, if $g(t) = c$, where c is a constant, then $u'' - au' + bcu = 0$ can be solved. Indeed, let $u = e^{zt}$, then substitution into the differential equation (C3) yields

$$(z^2 - az + bc)e^{zt} = 0,$$

which is satisfied for $z_1 = \frac{1}{2}(a + \sqrt{a^2 - 4bc})$ and $z_2 = \frac{1}{2}(a - \sqrt{a^2 - 4bc})$. If $bc = z_1 z_2 > 0$, then both z_1 and z_2 have the same sign. The general solution of (C3) is

$$u = u_1 e^{z_1 t} + u_2 e^{z_2 t}.$$

The corresponding solution (C4) for y becomes

$$y(t) = \frac{1}{b}\frac{z_1 + \frac{u_2}{u_1} z_2 e^{(z_2 - z_1)t}}{1 + \frac{u_2}{u_1} e^{(z_2 - z_1)t}}, \tag{C5}$$

where $z_2 - z_1 = -\sqrt{a^2 - 4bc}$. The initial condition $y(0) = y_0$ determines $\frac{u_2}{u_1} = -\frac{by_0 - z_1}{by_0 - z_2}$, so that, with $bc = z_1 z_2$,

$$y(t) = \frac{z_1 y_0 - c - (z_2 y_0 - c)e^{-t\sqrt{a^2 - 4bc}}}{by_0 - z_2 - (by_0 - z_1)e^{-t\sqrt{a^2 - 4bc}}}.$$

After introducing the expressions for z_1 and z_2, we find, in terms [39] of the hyperbolic tangent $\tanh(x) = \frac{e^{2x} - 1}{e^{2x} + 1}$,

$$y(t) = y_0 \frac{1 + \frac{(a - \frac{2c}{y_0})}{\sqrt{a^2 - 4bc}} \tanh\frac{t}{2}\sqrt{a^2 - 4bc}}{1 + \frac{(2by_0 - a)}{\sqrt{a^2 - 4bc}} \tanh\frac{t}{2}\sqrt{a^2 - 4bc}},$$

which can be further simplified with $\tanh(x + y) = \frac{\tanh(x) + \tanh(y)}{1 + \tanh(x)\tanh(y)}$ and defining

$$\Upsilon = \sqrt{a^2 - 4bc}$$

to

$$y(t) = \frac{a}{2b} + \frac{\Upsilon}{2b}\tanh\left[\frac{t}{2}\Upsilon + \operatorname{arctanh}\left(\frac{2by_0 - a}{\Upsilon}\right)\right]. \tag{C6}$$

An alternative and shorter derivation of (C6) when $g(t) = c$, follows from direct integration of the differential equation (C1):

$$\int_{y_0}^{y} \frac{du}{au - bu^2 - c} = t.$$

With $au - bu^2 - c = \frac{\Upsilon^2}{4b}\{1 - [\frac{2b}{\Upsilon}(u - \frac{a}{2b})]^2\}$, we obtain

$$t = \int_{y_0}^{y} \frac{du}{\frac{\Upsilon^2}{4b}\left\{1 - \left[\frac{2b}{\Upsilon}\left(u - \frac{a}{2b}\right)\right]^2\right\}} = \frac{2}{\Upsilon}\operatorname{arctanh}\frac{2b}{\Upsilon}\left(u - \frac{a}{2b}\right)\Big|_{y_0}^{y},$$

and inversion (i.e., solving for y) leads to (C6).

The derivative of (C6) is

$$\frac{dy(t)}{dt} = \frac{\Upsilon^2}{4b}\operatorname{sech}^2\left[\frac{t}{2}\Upsilon + \operatorname{arctanh}\left(\frac{2by_0 - a}{\Upsilon}\right)\right]. \tag{C7}$$

Suppose we define the metastable state when $\frac{dy(t)}{dt} \leqslant \varepsilon = 10^{-\gamma}$ (e.g., for $\gamma \leqslant 3$), then an estimate for the time to reach the metastable state is approximately

$$t_{\text{metastable}}(\varepsilon) = \frac{2}{\Upsilon}\left\{\operatorname{arccosh}\left(\sqrt{\frac{\Upsilon^2}{4b\varepsilon}}\right) - \operatorname{arctanh}\left(\frac{2by_0 - a}{\Upsilon}\right)\right\}.$$

If $a^2 - 4bc < 0$, then Υ is imaginary and (C6) becomes, with $\tanh(ix) = i\tan(x)$,

$$y(t) = \frac{a}{2b} + \frac{|\Upsilon|}{2b}\tan\left[-\frac{t}{2}|\Upsilon| + \arctan\left(\frac{2by_0 - a}{|\Upsilon|}\right)\right],$$

illustrating that poles occur at $t = \frac{-2}{\sqrt{4bc - a^2}}[\frac{\pi}{2} + k\pi - \arctan(\frac{2by_0 - a}{\sqrt{4bc - a^2}})]$ for $k \in \mathbb{Z}$, which is clearly not physical. Hence, we confine ourselves to the case where $\Upsilon = \sqrt{a^2 - 4bc}$ is real and positive. In order for $\sqrt{a^2 - 4bc}$ to be real, we require that $\frac{a^2}{4b} \geqslant c$. The maximum value, $c = \frac{a^2}{4b}$, in which case $u = u_1 e^{\frac{a}{2}t} + u_2 t e^{\frac{a}{2}t}$, leads to

$$y = \frac{1}{b}\frac{u'}{u} = \frac{a}{2b}\left(1 + \frac{\frac{2}{a}\frac{u_2}{u_1}}{1 + \frac{u_2}{u_1}t}\right),$$

which tends slowly to $y_\infty = \frac{a}{2b}$.

We investigate the impact of the real number c on $y(t)$, given that b and $\sqrt{a^2 - 4bc}$ are positive. The relations (C5) and (C6) show that $\lim_{t\to\infty} y(t) = y_\infty = \frac{z_1}{b}$. When $c = 0$, then $z_1 = a$ and $z_2 = 0$. If $c < 0$, then $\sqrt{a^2 - 4bc} > 0$ and both roots z_1 and z_2 are real but have opposite sign, $z_1 > 0 > z_2$, because $z_1 z_2 = bc < 0$. The more negative c is, the larger $\sqrt{a^2 - 4bc}$ and z_1, as well as the faster y tends to $y_\infty = \frac{z_1}{b}$. Since $y_\infty \leqslant 1$ and $y_\infty = \frac{z_1}{b}$, we find that $-c \leqslant b - a$. Combining the two constraints on c gives

$$a - b \leqslant c \leqslant \frac{a^2}{4b}. \tag{C8}$$

The derivative (C7) is positive if $\Upsilon^2 = a^2 - 4bc > 0$.

If $a < 0$ and $\sqrt{a^2 - 4bc} > 0$, then still $z_1 > z_2$. If $c > 0$, both z_1 and z_2 are negative, as well as $y_\infty = \frac{z_1}{b}$, which cannot

represent an average fraction of nodes. If $c < 0$, then $z_2 < 0$, but $z_1 > 0$ so that $y_\infty = \frac{z_1}{b}$ is physically possible.

APPENDIX D: PROOFS OF THEOREMS AND LEMMAS

1. Proof of Lemma 1

At an extremal time point $t^* = \theta$, where $\frac{dy}{dt^*}|_{t^*=\theta} = 0$, it follows from (B15) with $E[S^2] = y^2 + \text{var}[S]$ that

$$y^2(\theta) = \frac{1}{N}\sum_{k=1}^{N-1}(\tau\mu_k - 1)E\left[\zeta_k^2(\theta)\right] - \text{var}[S(\theta)].$$

Clearly, for $\tau < \frac{1}{\mu_1} < \tau_c$, the right-hand side is nonpositive, indicating that the only possible solution is $y(\theta) = 0$, corresponding to $\theta \to \infty$. We split the summation into two parts,

$$y^2(\theta) = \frac{1}{N}(\tau\mu_{N-1} - 1)E\left[\zeta_{N-1}^2(\theta)\right] + \frac{1}{N}\sum_{k=1}^{N-2}(\tau\mu_k - 1)E\left[\zeta_k^2(\theta)\right] - \text{var}[S(\theta)].$$

Introducing from (B7)

$$\zeta_{N-1}^2 = N(S - S^2) - \sum_{k=1}^{N-2}\zeta_k^2$$

yields after some manipulations and recognizing (11) of the correction R

$$\begin{aligned} y^2(\theta) - \left(1 - \frac{1}{\tau\mu_{N-1}}\right)y(\theta) + \text{var}[S(\theta)] &= \frac{1}{N}\sum_{k=1}^{N-2}\left(\frac{\mu_k}{\mu_{N-1}} - 1\right)E\left[\zeta_k^2(\theta)\right] \\ &= \frac{E[R(\theta)]}{N\mu_{N-1}} \geqslant 0. \end{aligned}$$

With the definition (10) of $\Psi(\theta) = \tau\mu_{N-1}\{\text{var}[S(\theta)] - \frac{E[R(\theta)]}{N\mu_{N-1}}\}$, we obtain the quadratic inequality

$$y^2(\theta) - \left(1 - \frac{1}{\tau\mu_{N-1}}\right)y(\theta) + \frac{\Psi(\theta)}{\tau\mu_{N-1}} = 0,$$

whose solution for $y(\theta)$ is (9). ■

2. Proof of Theorem 2

Introducing the lower bound (B17) into (4) yields

$$\frac{dy}{dt^*} \geqslant -y + \frac{\tau}{N}E\left[\sum_{j=1}^{N} d_j X_j\right] - \tau N\left(1 - \frac{1}{\omega}\right)E[S^2].$$

Using $\sum_{j=1}^{N} d_j X_j \geqslant d_{\min} NS$, $E[S^2] = (E[S])^2 + \text{var}[S]$, and $E[S] = y$, we end up with

$$\frac{dy}{dt^*} \geqslant (\tau d_{\min} - 1)y - N\tau\left(1 - \frac{1}{\omega}\right)y^2 - N\tau\left(1 - \frac{1}{\omega}\right)\text{var}[S], \tag{D1}$$

which is a Ricatti differential inequality associated to (C1). Appendix C demonstrates that, if

$$\max_{t\geqslant 0}\text{var}[S(t,\tau)] \leqslant \frac{(\tau d_{\min} - 1)^2}{4N^2\tau^2\left(1 - \frac{1}{\omega}\right)^2}$$

and $\tau > \frac{1}{d_{\min}}$, then a nonzero average fraction of infected nodes (after a sufficiently long time)

$$y_\infty \geqslant \frac{1}{2N\tau\left(1 - \frac{1}{\omega}\right)}((\tau d_{\min} - 1) + \sqrt{T})$$

with

$$T = (\tau d_{\min} - 1)^2 - 4N^2\tau^2\left(1 - \frac{1}{\omega}\right)^2\text{var}[S_\infty(\tau)].$$

The requirement for a positive discriminant T leads to (24), because then we are sure that $y_\infty > 0$, which implies that the epidemic threshold lies below this value, proving Theorem 2.

[1] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, Epidemic processes in complex networks, Rev. Mod. Phys. **87**, 925 (2015).

[2] T. M. Liggett, Improved upper bounds for the contract process critical value, Ann. Probab. **23**, 697 (1995).

[3] 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).

[4] P. Van Mieghem, F. D. Sahneh, and C. Scoglio, Exact Markovian SIR and SIS epidemics on networks and an upper bound for the epidemic threshold, in *Proceedings of the 53rd IEEE Conference on Decision and Control (CDC2014), December 15–17, Los Angeles*, CA (2014).

[5] P. Van Mieghem and R. van de Bovenkamp, Accuracy criterion for the mean-field approximation in SIS epidemics on networks, Phys. Rev. E **91**, 032812 (2015).

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

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

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

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

[10] P. Van Mieghem and R. van de Bovenkamp, Non-Markovian Infection Spread Dramatically Alters the SIS Epidemic Threshold in Networks, Phys. Rev. Lett. **110**, 108701 (2013).

[11] E. Cator, R. van de Bovenkamp, and P. Van Mieghem, Susceptible-Infected-Susceptible epidemics on networks with general infection and curing times, Phys. Rev. E **87**, 062816 (2013).

[12] M. Draief and L. Massoulié, *Epidemics and Rumours in Complex Networks*, London Mathematical Society Lecture Node Series, Vol. 369 (Cambridge University Press, Cambridge, 2010).

[13] P. Van Mieghem, Decay towards the overall-healthy state in SIS epidemics on networks, arXiv:1310.3980 (2013).
[14] R. van de Bovenkamp and P. Van Mieghem, Time to metastable state in SIS epidemics on graphs, in *Third International IEEE Workshop on Complex Networks and Their Applications, November 23–27, Marrakesh, Morocco* (2014).
[15] P. Van Mieghem, *Performance Analysis of Complex Networks and Systems* (Cambridge University Press, Cambridge, 2014).
[16] P. Van Mieghem, J. Omic, and R. E. Kooij, Virus spread in networks, IEEE/ACM Trans. Netw. **17**, 1 (2009).
[17] E. Cator and P. Van Mieghem, Second order mean-field SIS epidemic threshold, Phys. Rev. E **85**, 056111 (2012).
[18] P. Van Mieghem, D. Stevanović, F. A. Kuipers, C. Li, R. van de Bovenkamp, D. Liu, and H. Wang, Decreasing the spectral radius of a graph by link removals, Phys. Rev. E **84**, 016101 (2011).
[19] M. Fiedler, Algebraic connectivity of graphs, Czech. Math. J. **23**, 298 (1973).
[20] W. O. Kermack and A. G. McKendrick, A contribution to the mathematical theory of epidemics, Proc. R. Soc. London A **115**, 700 (1927).
[21] F. R. Gantmacher, *The Theory of Matrices*, Vol. II (Chelsea Publishing Company, New York, 1959).
[22] P. Van Mieghem, *Graph Spectra for Complex Networks* (Cambridge University Press, Cambridge, 2011).
[23] P. Van Mieghem, Interconnectivity structure of a general interdependent network, Phys. Rev. E **93**, 042305 (2016).
[24] P. Van Mieghem, The N-Intertwined SIS epidemic network model, Computing **93**, 147 (2011).
[25] E. Cator and P. Van Mieghem, Nodal infection in Markovian SIS and SIR epidemics on networks are non-negatively correlated, Phys. Rev. E **89**, 052802 (2014).
[26] A. Ganesh, L. Massoulié, and D. Towsley, The effect of network topology on the spread of epidemics, in *IEEE INFOCOM2005* (2005).
[27] J. Ricardo and G. de Mendonca, Precise critical exponents for the basis contact process, J. Phys. A **32**, L467 (1999).
[28] R. van de Bovenkamp and P. Van Mieghem, Survival time of the susceptible-infected-susceptible infection process on a graph, Phys. Rev. E **92**, 032806 (2015).
[29] D. Stevanović, *Spectral Radius of Graphs* (Academic Press, London, 2015).
[30] P. Van Mieghem and E. Cator, Epidemics in networks with nodal self-infections and the epidemic threshold, Phys. Rev. E **86**, 016116 (2012).
[31] T. Motzkin and E. G. Straus, Maxima for graphs and a new proof of a theorem of Turán, Can. J. Math. **17**, 533 (1965).
[32] T. M. Liggett, Multiple transition points for the contact process on the binary tree, Ann. Probab. **24**, 1675 (1996).
[33] N. Nagy and P. L. Simon, Monte Carlo simulation and analytic approximation of epidemic processes on large networks, Central Eur. J. Math. **11**, 800 (2013).
[34] P. Neal, The SIS great circle epidemic model, J. Appl. Probab. **45**, 513 (2008).
[35] H. S. Wilf, Spectral bounds for the clique and independence numbers of graphs, J. Comb. Theory, Ser. B **40**, 113 (1986).
[36] M. Boguñá, C. Castellano, and R. Pastor-Satorras, The Nature of the SIS Epidemic Threshold in Networks, Phys. Rev. Lett. **111**, 068701 (2013).
[37] P. Van Mieghem, Exact Markovian SIR and SIS epidemics on networks and an upper bound for the epidemic threshold, arXiv:1402.1731 (2014).
[38] P. Van Mieghem, SIS epidemics with time-dependent rates describing ageing of information spread and mutation of pathogens, Delft University of Technology, Report 20140615, www.nas.ewi.tudelft.nl/people/Piet/TUDelftReports, (2014).
[39] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover Publications, New York, 1968).