Autocorrelation of the susceptible-infected-susceptible process on networks

Qiang Liu[*] and Piet Van Mieghem[†]

Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, 2600 GA Delft, The Netherlands

In this paper, we focus on the autocorrelation of the susceptible-infected-susceptible (SIS) process on networks. The N-intertwined mean-field approximation (NIMFA) is applied to calculate the autocorrelation properties of the exact SIS process. We derive the autocorrelation of the infection state of each node and the fraction of infected nodes both in the steady and transient states as functions of the infection probabilities of nodes. Moreover, we show that the autocorrelation can be used to estimate the infection and curing rates of the SIS process. The theoretical results are compared with the simulation of the exact SIS process. Our work fully utilizes the potential of the mean-field method and shows that NIMFA can indeed capture the autocorrelation properties of the exact SIS process.

I. INTRODUCTION

The susceptible-infected-susceptible (SIS) process [1] is a basic epidemic model which models the spread of viruses, information, opinions, and computer malware on networks. In the SIS model, each node in the network can be either infected or susceptible (healthy). The infection state of node j for $j = 1, \ldots, N$ at time t is denoted by a Bernoulli random variable $X_j(t)$: infected $X_j(t) = 1$ or susceptible (healthy) $X_j(t) = 0$. The SIS model has simple local rules that nodes can be infected by their infected neighbors and be cured by themselves. The infection and curing processes are independent and both Poisson processes with infection rate β and curing rate δ, respectively. By tuning the effective infection rate $\tau \triangleq \beta/\delta$, the phase transition of infection persistence emerges at an epidemic threshold determined by the network [2,3]. If the effective infection rate τ is below the epidemic threshold, then the virus dies out quickly and every node becomes healthy. Above the threshold, the infection can persist in the network for a very long time [4].

In this paper, we focus on the autocorrelation of the SIS process. Locally, an individual node in the network can be infected and cured repeatedly so that the infection state $X_j(t)$ at two different time points can be autocorrelated. The autocorrelation of the infection state of a node j between time s and t is

$$\rho_j(s,t) \triangleq \frac{E[X_j(s)X_j(t)] - E[X_j(s)]E[X_j(t)]}{\sqrt{\text{Var}[X_j(s)]\text{Var}[X_j(t)]}}. \quad (1)$$

The numerator on the right-hand side in (1) is the covariance of the infection state $X_j(s)$ and $X_j(t)$, and the denominator normalizes the covariance. If time $t = s$, then the infection state $X_j(s)$ is fully correlated with itself, and the autocorrelation is $\rho_j(s,s) = 1$. If $X_j(s)$ and $X_j(t)$ are independent, then the autocorrelation is $\rho_j(s,t) = 0$. The autocorrelation is

symmetric: $\rho_j(s,t) = \rho_j(t,s)$. Given the initial infection state of the network $[X_1(0), \ldots, X_N(0)]$, the infection states $X_j(s)$ and $X_j(t)$ are positively correlated [5], Corollary 1 such that $E[X_j(s)X_j(t)] \geqslant E[X_j(s)]E[X_j(t)]$ and the autocorrelation $\rho_j(s,t) \geqslant 0$. The autocorrelation $\rho_j(s,t)$ contains one second-moment term $E[X_j(s)X_j(t)]$, but the rest of the terms can be calculated given the first-moment infection probabilities $E[X_j(s)]$ and $E[X_j(t)]$.

The autocorrelation contains information about the change of the infection state of each node. A large autocorrelation implies that the change of the infection state is slow, and the infection state is more likely to be identical between time s and t. While a smaller autocorrelation indicates that the infection state between time s and t is more independent. Globally, the fluctuating fraction of infected nodes $I(t) \triangleq \frac{1}{N} \sum_{j=1}^{N} X_j(t)$ is also autocorrelated, and the autocorrelation and its spectral analysis of $I(t)$ in real epidemics can be traced back to Anderson *et al.* [6]. By analyzing the autocorrelation and its spectrum of the incidence data of pertussis, mumps, and measles, Anderson *et al.* [6] indicate statistically significant seasonal and the longer-term resurgence of those diseases and find that vaccination increases the periods of the longer-term oscillations of the incidence data. However, in the basic networked SIS model, the autocorrelation of the infection state is infeasible to be calculated, because the SIS model is a 2^N-state Markov process [3,7] and the computational complexity is exponentially high regarding of the network size N. Previously, Meier *et al.* [8], Supplementary Information E analyzed the correlation of the infection state of the SIS model for small time intervals, but the calculation involves higher-order moments. In this paper, we apply the N-intertwined mean-field approximation (NIMFA) [3] to study the autocorrelation of the infection state $X_j(t)$ and the fraction of infected nodes $I(t)$ both in the transient and steady states. Particularly in the steady state, we derive the explicit formula of the autocorrelation of the infection state, which is an exponentially decreasing function of time delay. The accuracy of the NIMFA autocorrelation is evaluated by simulating the exact SIS process. The result indicates that NIMFA, as an approximate stochastic process,

---
[*]Q.L.Liu@tudelft.nl
[†]P.F.A.VanMieghem@TUDelft.nl

well captures the autocorrelation properties of the exact SIS process. Moreover, the autocorrelation can also be used to estimate the infection and curing rates of the SIS process.

II. THE SIS PROCESS ON NETWORKS

The undirected and unweighted network with N nodes is denoted by its $N \times N$ symmetric adjacency matrix A: If node i and j are connected, then $a_{ij} = a_{ji} = 1$; otherwise, $a_{ij} = a_{ji} = 0$. The infection probability of node j at time t is just the expectation of its infection state $E[X_j(t)]$, and the prevalence is defined as the expectation of the fraction of infected nodes $y(t) \triangleq E[I(t)]$.

A. The exact SIS process

The SIS process can be described by a Markov process, and there are 2^N states in total, including one all-healthy absorbing state [3,7]. The state transition of each node in the 2^N-state Markov process can be described as

$$X_j(t) : 0 \rightarrow 1 \text{ with rate: } \beta \sum_{i=1}^{N} a_{ji} X_i(t)$$

$$X_j(t) : 1 \rightarrow 0 \text{ with rate: } \delta. \qquad (2)$$

Since the all-healthy state is absorbing and the network is finite, the SIS process will enter the absorbing state when $t \rightarrow \infty$. However, the SIS process can also stay in the metastable state for a long time where the infection probability of every node is almost constant.

The infection probability of node j for $j = 1, \ldots, N$ follows the governing equation [9],

$$\frac{dE[X_j(t)]}{dt} = -\delta E[X_j(t)] + \beta \sum_{i=1}^{N} a_{ji} E[X_i(t)]$$

$$- \beta \sum_{i=1}^{N} a_{ji} E[X_j(t)X_i(t)]. \qquad (3)$$

Equation (3) describes the exact Markovian SIS process, but higher-order moments of the infection states $E[X_j(t)X_i(t)]$ are involved in Eq. (3). In total, $2^N - 1$ equations are needed to solve the process [9, p. 452] and the complexity increases exponentially with network size N. Furthermore, the analysis of the SIS process is not tractable without approximation, not even for the complete graph [10].

B. The N-intertwined mean-field approximation

NIMFA [11] approximates the exact Markovian SIS process by assuming independence $E[X_i(t)X_j(t)] = E[X_i(t)]E[X_j(t)]$, which is equivalent to approximating the infection rate due to all neighbors $\beta \sum_{i=1}^{N} a_{ji} X_i(t)$ in (2) by its mean $\beta \sum_{i=1}^{N} a_{ji} E[X_i(t)]$. For Bernoulli random variables, uncorrelation $E[X_i(t)X_j(t)] = E[X_i(t)]E[X_j(t)]$ and independence $\Pr[X_i(t), X_j(t)] = \Pr[X_i(t)]\Pr[X_j(t)]$ are equivalent [12, footnote 5]. Under NIMFA, the governing

equation is

$$\frac{dv_j(t)}{dt} = -\delta v_j(t) + \beta(1 - v_j(t)) \sum_{j=1}^{N} a_{ji} v_i(t), \qquad (4)$$

where $v_j(t)$ is the NIMFA infection probability of node j at time t and $v_j(t)$ approximates the exact infection probability $E[X_j(t)]$. The NIMFA epidemic threshold is $\tau_c^{(1)} \triangleq \frac{1}{\lambda_1}$, where λ_1 is the largest eigenvalue of the adjacency matrix A. If the effective infection rate $\tau \triangleq \frac{\beta}{\delta} > \tau_c^{(1)}$, then the infection can persist on the network and the steady-state infection probability $v_{j\infty} \triangleq \lim_{t \to \infty} v_j(t) > 0$ is constant [3,13]. The steady state of NIMFA corresponds to the metastable state of the exact SIS process. If $\tau < \tau_c^{(1)}$, then the NIMFA SIS process will eventually enter the all-healthy state $v_{j\infty} = 0$. NIMFA has been successfully applied to analyze the first-order moments of the SIS process [3]. For example, the NIMFA infection probability $v_j(t)$ and the prevalence $y^{(1)}(t) \triangleq \frac{1}{N} \sum_{j=1}^{N} v_j(t)$ well approximate the expectations of the infection state $X_j(t)$ and the fraction of infected nodes $I(t) \triangleq \frac{1}{N} \sum_{j=1}^{N} X_j(t)$, respectively. However, NIMFA has not yet been applied to approximate the autocorrelation properties. Since NIMFA omits the correlation between neighbors, the autocorrelation is the only second-moment property that is possibly captured by NIMFA.

To avoid ambiguity, we denote the NIMFA infection state of node j at time t by another Bernoulli random variable $V_j(t)$: infected $V_j(t) = 1$ and susceptible $V_j(t) = 0$. Thus, we actually approximate the statistical properties of the infection state $X_j(t)$ by those of $V_j(t)$ in NIMFA. In the steady state $t \to \infty$ and $dv_j(t)/dt = 0$ for $j = 1, \ldots, N$, we denote the infection state of node j by $V_{j\infty}(t) \triangleq \lim_{t \to \infty} V_j(t)$.

Under NIMFA, the transition of the infection state $V_j(t)$ of node j following Eq. (4) can be denoted by a two-state Markov process [14], and the transition rate of $V_j(t) : 0 \rightarrow 1$ becomes a determined function of time. The whole system is composed of N intertwined 2-state Markov processes instead of being a 2^N-state Markov process. Corresponding to (2), the transition of the NIMFA infection state $V_j(t)$ is

$$V_j(t) : 0 \rightarrow 1 \text{ with rate: } \tilde{\beta}_j(t) \triangleq \beta \sum_{j=1}^{N} a_{ji} v_i(t)$$

$$V_j(t) : 1 \rightarrow 0 \text{ with rate: } \delta. \qquad (5)$$

The infinitesimal generator of the Markov process (5) is

$$Q_j(t) \triangleq \begin{bmatrix} -\tilde{\beta}_j(t) & \tilde{\beta}_j(t) \\ \delta & -\delta \end{bmatrix}. \qquad (6)$$

III. AUTOCORRELATION IN THE STEADY STATE

In the steady state, the NIMFA autocorrelation of the infection state of node j with time lag h is defined by

$$R_{j\infty}(h) \triangleq \frac{E[V_{j\infty}(t)V_{j\infty}(t+h)] - v_{j\infty}^2}{\text{Var}[V_{j\infty}(t)]}, \qquad (7)$$

where $\text{Var}[V_{j\infty}(t)] = v_{j\infty} - v_{j\infty}^2$ since $V_{j\infty}(t) \in \{0,1\}$ is a Bernoulli random variable. By further derivation (see Appendix A), we obtain the autocorrelation as a function of the

steady-state infection probability $v_{j\infty}$ and the curing rate δ,

$$R_{j\infty}(h) = e^{-\frac{\delta}{1-v_{j\infty}}h}, \qquad (8)$$

where we assume that the time lag h is positive without loss of generality. Since the autocorrelation is symmetric $R_{j\infty}(h) = R_{j\infty}(-h)$, $R_{j\infty}(h) = e^{\frac{\delta}{1-v_{j\infty}}h}$ for $h < 0$. The NIMFA infection probability $v_{j\infty}$ in (8) can be obtained by solving the NIMFA Eq. (4) numerically.

With a fixed δ and time lag h, the autocorrelation $R_{j\infty}(h)$ in (8) decreases with the infection rate β because the infection probability $v_{j\infty}$ increases correspondingly. A larger infection rate β implies a faster state transition from healthy to infected, and the autocorrelation of the infection state is smaller consequently. A larger δ leads to a faster transition from infected to healthy, but, simultaneously, the infection probability of each neighbor becomes smaller. Therefore, the state transition of each node is slower from the healthy state to the infected state, and the effect of the curing rate δ is unclear. Only in special networks can the effect of the curing rate δ be determined. For example, the infection probabilities of all nodes are equal to $1 - \frac{1}{k\tau}$ in a k-regular graph [3], and then the autocorrelation function becomes

$$R_{j\infty;k\text{-regular}}(h) = e^{-\beta k h}. \qquad (9)$$

Formula (9) indicates that the autocorrelation of the infection state does not depend on the curing rate δ in regular graphs, which enables us to adjust the autocorrelation while keeping the effective infection rate τ unchanged. In regular graphs, the effect of the decrease (increase) of $v_{j\infty}$ is exactly compensated by the increase (decrease) of δ in (8). The autocorrelation under other mean-field approximations can also be derived with the same procedure. For example, the heterogeneous mean-field approximation (HMF) assumes statistical equivalence among the nodes with the same degree [2], and the autocorrelation under HMF has the same form as the NIMFA autocorrelation (see Appendix B). In the case of regular graphs, HMF is equivalent [15] to NIMFA and then their approximate autocorrelations are identical.

Generally, the NIMFA infection probability of node j with degree d_j for $j = 1, \dots, N$ is bounded by [3]

$$1 - \frac{1}{1 + \tau d_j - \frac{d_j}{d_{\min}}} \leqslant v_{j\infty} \leqslant 1 - \frac{1}{1 + \tau d_j}$$

in a connected network with minimum degree $d_{\min}$, and the NIMFA autocorrelation (8) is thus bounded by

$$e^{-(1+\tau d_j)\delta h} \leqslant R_{j\infty}(h) \leqslant e^{-\left[1 + d_j\left(\tau - \frac{1}{d_{\min}}\right)\right]\delta h}. \qquad (10)$$

The largest eigenvalue of the adjacency matrix λ_1 follows $\lambda_1 \geqslant d_{\min}$, and then the effective infection rate τ can either be larger or smaller than $1/d_{\min}$ when τ is above the threshold $\tau_c^{(1)} = \frac{1}{\lambda_1}$. Equation (8) indicates that the autocorrelation has another upper bound

$$R_{j\infty} < e^{-\delta h} \qquad (11)$$

when $v_{j\infty} > 0$ (i.e., above the threshold). If $\frac{1}{\lambda_1} < \tau < \frac{1}{d_{\min}}$, then

$$e^{-\delta h} < e^{-\left[1 + d_j\left(\tau - \frac{1}{d_{\min}}\right)\right]\delta h}$$

and the upper bound (11) is tighter. If $\tau > \frac{1}{d_{\min}}$, then the upper bound in (10) is tighter, and we can rewrite (10) as

$$e^{-(1+\tau d_j)\delta h} \leqslant R_{j\infty}(h) \leqslant e^{-(1+\tau d_j)\delta h} e^{\left(\frac{d_j}{d_{\min}}\right)\delta h}. \qquad (12)$$

In (12), the upper bound is just the product of the lower bound and the term $e^{(d_j/d_{\min})\delta h} > 1$. In a network with large degree deviation $d_j/d_{\min}$, the bound (12) is loose. In the regular graph, $\lambda_1 = d_{\min}$, and the upper bound achieves the exact NIMFA autocorrelation (9) while the lower bound does not.

In a heterogeneous network, e.g., the scale-free network, the degree d_j can diverge in the thermodynamic limit $N \to \infty$. Thus, if $\tau > \frac{1}{d_{\min}}$ and $d_j \to \infty$, then both the upper and lower bound in (12) converge to zero, and the autocorrelation $R_{j\infty}(h) = 0$. If $\frac{1}{\lambda_1} < \tau < \frac{1}{d_{\min}}$ and $d_j \to \infty$, then the lower bound $R_{j\infty}(h) > e^{-(1+\tau d_j)\delta h}$ converges to zero. Consequently, the autocorrelation is loosely bounded by $0 \leqslant R_{j\infty}(h) \leqslant \exp(-\delta h)$.

From a global point of view, the fraction of infected nodes $I(t) = \frac{1}{N}\sum_{j=1}^{N} X_j(t)$ in the steady state can be approximated by $I_\infty^{(1)}(t) \triangleq \frac{1}{N}\sum_{j=1}^{N} V_{j\infty}(t)$. The autocorrelation of $I_\infty^{(1)}(t)$ is just a linear combination of the autocorrelation of each node (see Appendix A),

$$R_{I^{(1)}\infty}(h) = \frac{\sum_{j=1}^{N}\left(v_{j\infty} - v_{j\infty}^2\right)R_{j\infty}(h)}{\sum_{j=1}^{N}\left(v_{j\infty} - v_{j\infty}^2\right)}. \qquad (13)$$

IV. AUTOCORRELATION IN THE TRANSIENT STATE

In this section, we consider the NIMFA autocorrelation of the SIS process at two arbitrary time points s and t, respectively. Different from that in the steady state in Sec. III, the infinitesimal generator (6) is a determined function of time given the initial state. The two-state Markov process (5) of each node is thus a time-inhomogeneous process. Calculating the process (5) allows us to analyze the autocorrelation of the epidemic process in the transient regime before the metastable state or the regime before the all-healthy steady state when the effective infection rate $\tau < \tau_c^{(1)}$.

We denote the NIMFA autocorrelation of node j between time s and t as

$$R_j(s,t) \triangleq \frac{E[V_j(s)V_j(t)] - v_j(s)v_j(t)}{\sqrt{\left[v_j(s) - v_j^2(s)\right]\left[v_j(t) - v_j^2(t)\right]}}. \qquad (14)$$

Following a similar derivation as Eq. (13) in the steady state, the autocorrelation of the fraction of infected nodes $R_{I^{(1)}}(s,t)$ is also a linear combination of the autocorrelation of each node,

$$R_{I^{(1)}}(s,t) = \frac{\sum_{j=1}^{N}\sqrt{\left[v_j(s) - v_j^2(s)\right]\left[v_j(t) - v_j^2(t)\right]}R_j(s,t)}{\sqrt{\sum_{j=1}^{N}\left[v_j(s) - v_j^2(s)\right]\sum_{j=1}^{N}\left[v_j(t) - v_j^2(t)\right]}}. \qquad (15)$$

Similarly to the steady-state autocorrelation in Sec. III, we only use the infection probabilities in the calculation, and the joint expectation $E[V_j(s)V_j(t)]$ in (14) becomes a crucial term. The calculation of the joint expectation $E[V_j(s)V_j(t)] = v_j(s)\Pr[V_j(t) = 1 | V_j(s) = 1]$ involves the 2×2 time-dependent transition matrix $P_j(s,t)$ of which the

element is $[P_j(s,t)]_{kl} = \Pr[V_j(t) = l - 1 | V_j(s) = k - 1]$. The computation of the autocorrelation functions (14) and (15), requires us first to calculate the matrix $P_j(s,t)$.

The matrix $P_j(s,t)$ follows the time-inhomogeneous Kolmogorov forward equation

$$\frac{dP_j(s,t)}{dt} = P_j(s,t)Q_j(t), \tag{16}$$

where $Q_j(t)$ is the NIMFA infinitesimal generator (6). We can apply the Magnus expansion [16,17] to analyze the NIMFA transition matrix $P_j(s,t)$ in Eq. (16). A brief introduction of the Magnus expansion can be found in Appendix C. Although the calculation of the exact NIMFA transition probability $P_j(s,t)$ is not possible, approximations of $P_j(s,t)$ allowing a fair comparison between NIMFA and the exact SIS process can be made with restricted error. First, there exists a 2×2 matrix $\Omega(s,t;j)$ such that the solution of Eq. (16) is $P_j(s,t) = \exp[\Omega(s,t;j)]$. Second, if (see the derivation of (C5) for details in Appendix C)

$$0 < t - s < T \triangleq \frac{\pi}{\sqrt{\beta^2 d_j^2 + \delta^2}}, \tag{17}$$

then the exponent matrix $\Omega(s,t;j)$ can be expanded into a convergent Magnus series $\Omega(s,t;j) = \sum_{k=1}^{\infty} \Omega_k(s,t;j)$. Specifically, by only preserving the first term, i.e., $\Omega_1(s,t;j) = \int_s^{s+h} Q_j(t)dt$, in the convergent Magnus series of $\Omega(s,t;j)$, we can achieve a third-order accuracy (see Appendix C) for the time length $h = t - s$, i.e.,

$$P_j(s,s+h) = \exp\left[\int_s^{s+h} Q_j(t)dt\right] + O(h^3). \tag{18}$$

Equation (18) holds because $\exp[X + O(h^k)] = \exp(X) + O(h^k)$ holds for a matrix X as can be verified by evaluating their power series. Using the Taylor expansion of the infinitesimal generator $Q_j(t) = \sum_{k=0}^{\infty} \frac{1}{k!} \frac{d^k Q_k(u)}{du^k}\big|_{u=s}(t-s)^k$ at time s, the solution (18) becomes

$$P_j(s,s+h) = \exp\left[Q_j(s)h + \frac{dQ_j(t)}{dt}\bigg|_{t=s} \frac{h^2}{2}\right] + O(h^3). \tag{19}$$

Only the first two terms of the Taylor expansion of the infinitesimal generator $Q_j(t)$ are preserved in (19) since the error is $O(h^3)$ in (18). The first term on the right-hand side of (19) can be calculated by matrix diagonalization described in Appendix A. The derivative of the infinitesimal generator $dQ_j(t)/dt$ involves $\frac{d\tilde{\beta}_j(t)}{dt}$ from Eq. (6), which is

$$\tilde{\beta}_j'(t) \triangleq \frac{d\tilde{\beta}_j(t)}{dt} = \beta \sum_{i \in \mathcal{N}_j} \frac{dv_i(t)}{dt}$$

$$= \beta \sum_{i \in \mathcal{N}_j} \left\{ -\delta v_i(t) + \beta[1 - v_i(t)] \sum_{k \in \mathcal{N}_i} v_k(t) \right\}, \tag{20}$$

where $\mathcal{N}_j$ denotes the neighbors of node j. The calculation in Eq. (20) involves the infection probabilities of two-hop

neighbors of node j. Specifically, the transition probability that node j remains infected after h time units is

$$\Pr[V_j(s+h) = 1 | V_j(s) = 1] = [P_j(s,s+h)]_{22}$$

$$= \frac{2\delta e^{-(\tilde{\beta}_j(s)+\delta)h - \tilde{\beta}_j'(s)h^2/2} + 2\tilde{\beta}_j(s) + \tilde{\beta}_j'(s)h}{2\tilde{\beta}_j(s) + 2\delta + \tilde{\beta}_j'(s)h}$$

$$+ O(h^3).$$

Different from that in the steady state [see Eq. (8)], the infection probabilities of neighbors of node j always appear in the calculation of the transition matrix $P_j(s,t)$ in the transient state as indicated in (20). Higher-order accuracy is also possible by preserving more terms of the Magnus series, and higher-order derivative $d^k \tilde{\beta}_j(t)/dt$, which can be calculated by the infection probabilities of all nodes within $k+1$ hops from node j, is involved. For example, if we preserve the second term in the Magnus expansion of $\Omega(s,s+h;j)$, which can be calculated by the Taylor expansion as

$$\Omega_2(s,s+h;j) = \delta \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \int_s^{s+h} \int_s^{t_1} [\tilde{\beta}_j(t_1) - \tilde{\beta}_j(t_2)] dt_2 dt_1$$

$$= \delta \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \left[\frac{1}{6} \frac{d\tilde{\beta}_j(t)}{dt}\bigg|_{t=s} h^3 \right.$$

$$\left. + \frac{1}{12} \frac{d^2 \tilde{\beta}_j(t)}{dt^2}\bigg|_{t=s} h^4 \right] + O(h^5),$$

then we can achieve an accuracy of $O(h^5)$ because (see Appendix C)

$$P_j(s,s+h) = \exp\left[\sum_{i=1}^{2} \Omega_i(s,s+h;j)\right] + O(h^5)$$

and the calculation involves the infection probabilities of neighbors within three hops. For NIMFA, preserving more terms is not always reasonable, because the infection probability of each node can only be solved numerically. When more Magnus terms are preserved, the inaccuracy is mainly caused by the numerical method which solves the nonlinear NIMFA Eq. (4). For example, using the fourth-order Runge-Kutta method [17, p. 200], the error of the infection probabilities is of order $O(h^4)$.

For a time interval $t - s > T$, the Magnus expansion of the exponent $\Omega(s,t;j)$ may not converge. The time interval (s,t) can be divided into subintervals with length $h < T$ in which the Magnus series converges. The NIMFA transition matrix between time s and t can be written as

$$P_j(s,t) = \prod_{k=1}^{(t-s)/h} P_j(s+(k-1)h, s+kh) \tag{21}$$

by the Chapman-Kolmogorov equation [see Eq. (C2)]. Equation (21) is also applicable to a small time interval $t - s < T$ to obtain a more accurate result. An rth-order accuracy regarding the time delay h is achieved for the transition matrix $P_j(s,t)$ using Eq. (21) if the accuracy is $O(h^{r+1})$ for each $P_j(s+(k-1)h, s+kh)$.
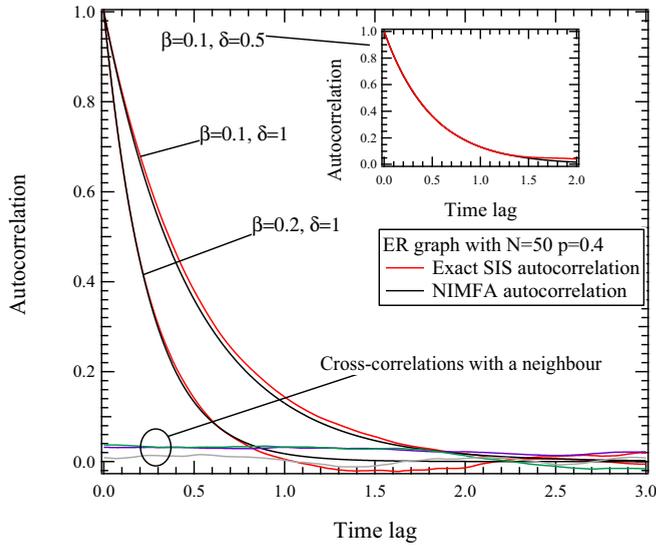
FIG. 1. A randomly selected node is evaluated in an Erdős-Rényi (ER) network with the link connecting probability 0.4 and $N = 50$. The autocorrelation is approximately constant for different value of δ. The cross correlation between the node with one of its neighbors is also plotted, which is almost zero.

The analysis in this section allows us to calculate and compare the NIMFA autocorrelation with the exact SIS process since the error can be controlled, even though the exact NIMFA autocorrelation is not feasible in the transient state.

V. NUMERICAL AND SIMULATION RESULTS

In this section, we compare the NIMFA autocorrelation with the autocorrelation of the exact SIS process from the simulation. The simulation of the exact SIS process is implemented by the Gillespie algorithm (Monte Carlo method) [18–20] and the theoretical results are obtained by solving the NIMFA Eq. (4) numerically (fourth-order Runge-Kutta method [17, p. 200]). In the steady state, we run the simulation for 40 000 time units with the curing rate $\delta = 1$ and sample the infection state of each node every 0.001 time unit. In other words, we obtain the infection state $X_j(n/1000)$ for $n = 0, 1, \ldots, 4 \times 10^7$ from simulation. We only use the state sequence sampled after $t = 10\,000$ to ensure that the SIS process is in the metastable state. Moreover, the time series of the fraction of infected nodes can be calculated as $I(n/1000) = \frac{1}{N} \sum_{j=1}^{N} X_j(n/1000)$. In the transient state, 10^4 realizations of the infection states $X_j(s)$ and $X_j(t)$ are obtained to calculate the autocorrelation between two arbitrary time s and t.

A. Steady state

Figures 1 to 3 show the NIMFA autocorrelation and the simulated autocorrelation of the infection state of randomly selected nodes in an Erdős-Rényi (ER) graph, a regular graph with degree 26, and a star graph, respectively. The NIMFA autocorrelation $R_{j\infty}(h)$ is a very accurate approximation on those graphs. Figure 1 shows that the autocorrelation of the infection state is not sensitive to the value of the curing rate δ, which is reasonable because the deviation of the degree is small
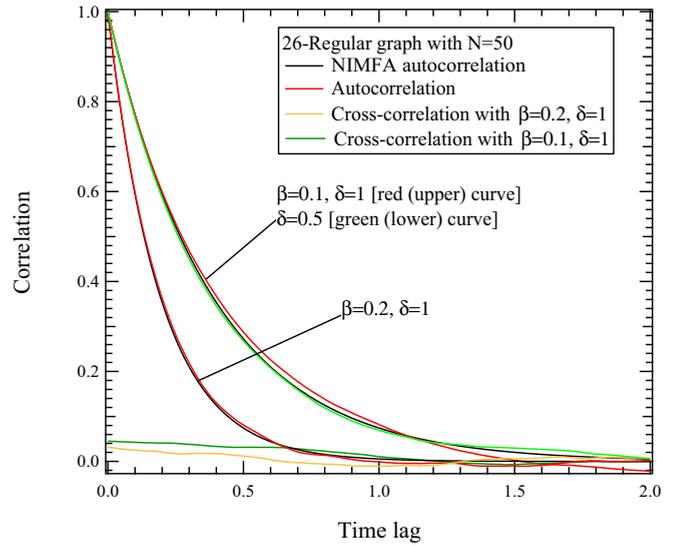
FIG. 2. The autocorrelation of the infection state of a 26-regular graph with $N = 50$. The results are similar to those of the ER graph. The autocorrelation is invariant to δ.

and the result is similar to that of the regular graph in Fig. 2. In Fig. 2, the autocorrelation of the infection state is identical to formula (9) that the autocorrelation is invariant to the curing rate δ in regular graphs. Figure 3 shows the autocorrelation of the infection state in a star graph. The autocorrelation of the hub node is much smaller than that of the leaf nodes since the infection probability of the hub node is larger. The cross correlation of the infection states between neighbors shown in Fig. 1 to 3 is approximately 0, which leads to the effectiveness of NIMFA since NIMFA omits the cross correlation between neighbors.
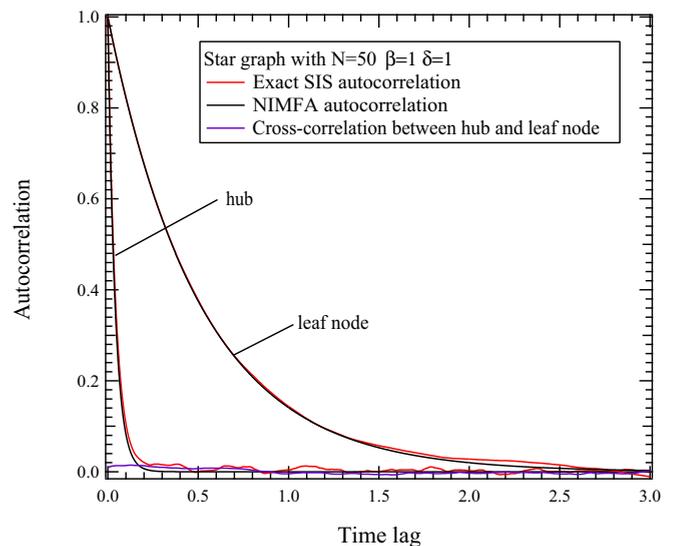
FIG. 3. The autocorrelation of the infection state of the hub and a leaf node in a star graph with $N = 50$. The NIMFA autocorrelation shows a very good approximation and the cross correlation between hub and leaf nodes is approximately 0.
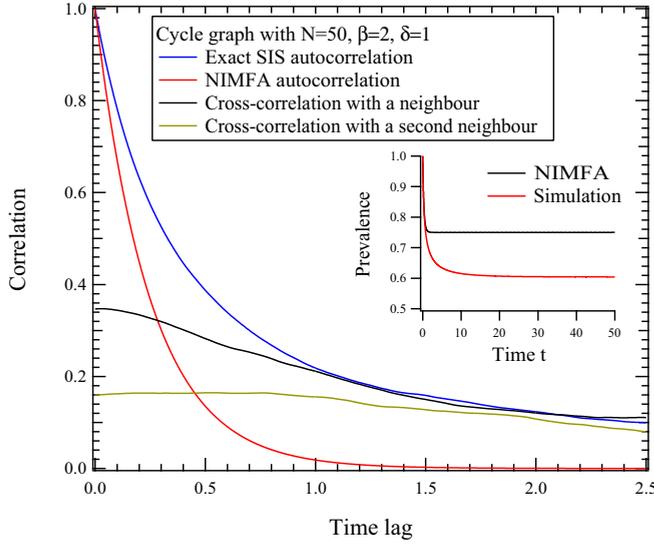
FIG. 4. The correlation of the infection state of a node and the prevalence in a cycle graph with $N = 50$. Initially all nodes are infected to prevent the inaccuracy caused by early die-out [24]. The NIMFA autocorrelation is much smaller than the exact one, and the cross correlations between neighbors and second-hop neighbors are very large.

Figure 4 shows the autocorrelation of the infection state of a node in a cycle graph and NIMFA fails to capture the autocorrelation. Actually, NIMFA also fails to approximate the prevalence as shown in Fig. 4. In the situation of the cycle graph, the cross correlation of the infection states between neighbors is much larger than zero and NIMFA itself is a bad approximation. The accuracy of mean-field methods has been studied in Refs. [21–23], which is beyond the scope of this paper.

We also calculate the autocorrelation of the fraction of infected nodes $R_{I^{(1)}}(h)$. Figure 5 shows that NIMFA can also
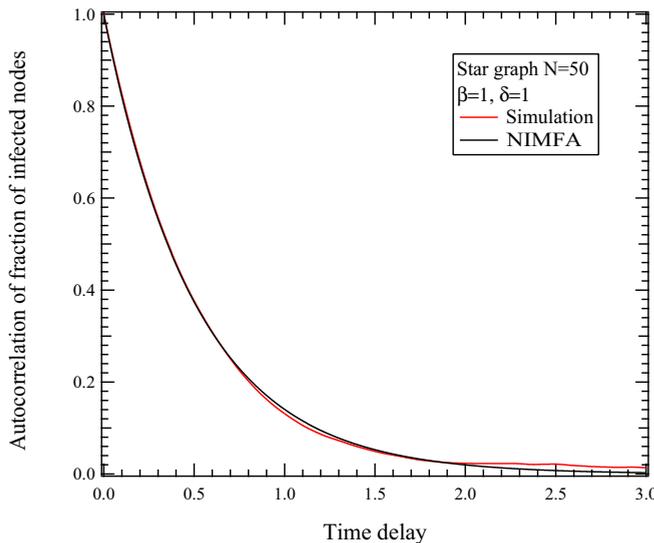
FIG. 5. The autocorrelation of the fraction of infected nodes $I(t)$ in the metastable state.
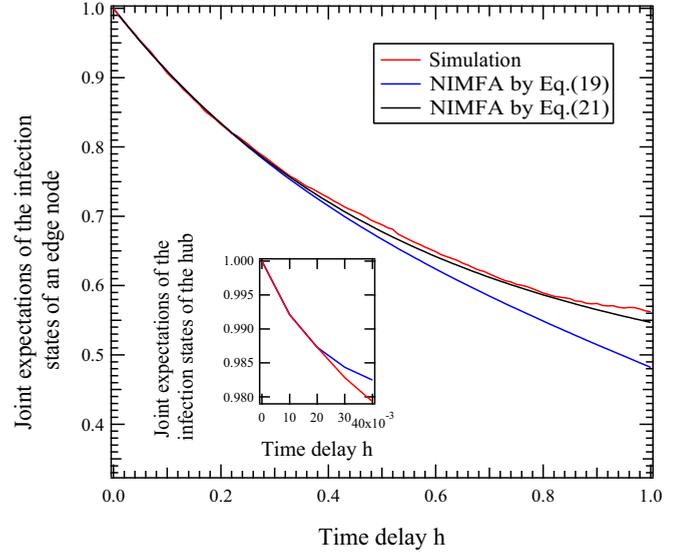
FIG. 6. The joint expectation of the infection state $E[X_j(0)X_j(h)]$ and the corresponding NIMFA approximation $E[V_j(0)V_j(h)]$ of the SIS process on the star graph.

approximate the autocorrelation of the fraction of infected nodes in the star graph corresponding to Fig. 3.

B. Transient state

In the transient state, we validate the NIMFA autocorrelation on the star graph where the NIMFA infection probabilities are accurate while nodes have very different degrees.

Figure 6 shows the joint expectation of the infection states $E[X_j(0)X_j(h)]$ and the corresponding NIMFA approximation $E[V_j(0)V_j(h)]$ of the leaf and hub nodes. For the leaf node and the hub node, the convergent time delay h of the Magnus series of $\Omega(s, s + h)$ are $h < T \approx 2.221$ and $h < T \approx 0.064$ from (17), respectively. Figure 6 indicates that the NIMFA joint expectation $E[V_j(0)V_j(h)]$ (the blue lower curve) is accurate comparing with the exact joint expectation $E[X_j(0)X_j(h)]$ for a small time delay h, i.e., $h < 0.2$ for the leaf node. For a large time delay, the inaccuracy is due to either the omission of term $O(h^3)$ in (19) or that the NIMFA transition probability matrix $P_j(s, t)$ itself is a bad approximation, but we can eliminate the possibility of the latter using Eq. (21). As the black middle curve in Fig. 6 indicated, the NIMFA joint expectation $E[V_j(0)V_j(h)]$ is indeed a good approximation using Eq. (21) with subinterval length 0.01.

From a global point of view of the network, Fig. 7 presents the autocorrelation of the fraction of infected nodes $R_I(0.5, 0.5 + h)$ and the corresponding NIMFA approximation $R_{I^{(1)}}(0.5, 0.5 + h)$, which are in the transient state of the SIS process before the metastable state. The exact autocorrelation is well fitted by NIMFA. Interestingly, the decay of the autocorrelation in the transient state is also exponential as shown in Fig. 7, but we cannot demonstrate exponential decay as opposed to the steady state.

In this section, we have tested our method on different networks with size 50, but for larger networks, the results are similar. In a conclusion, NIMFA captures the autocorrelation properties of the exact SIS process except in the cases that
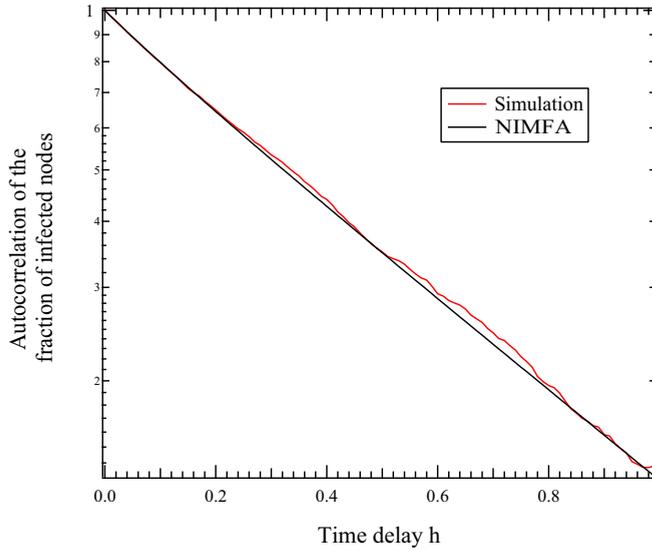
FIG. 7. The autocorrelation of the fraction of infected nodes $R_I(0.5, 0.5 + h)$ and the corresponding NIMFA approximation $R_{I^{(1)}}(0.5, 0.5 + h)$ of the SIS process on the star graph.

NIMFA is not applicable even for approximating the first-moment properties, i.e., the infection probabilities $E[X_j(t)]$ and the prevalence $y(t)$.

VI. ESTIMATING THE CURING RATE δ AND THE INFECTION RATE β: AN APPLICATION

In real epidemics, a disease agency may have the infection-state data by monitoring individuals periodically but no information about the rates. We consider the reverse problem of estimating the curing rate δ and the infection rate β, given the sequence $X_j(t + \Delta), X_j(t + 2\Delta), \ldots, X_j(t + n\Delta)$ of the infection state of node j in the metastable state. From Eq. (8), the curing rate is

$$\delta = -(1 - v_{j\infty}) \frac{\ln[R_{j\infty}(h)]}{h}. \qquad (22)$$

Formula (22) can be used to estimate the curing rate δ of the SIS process. In formula (22), we can approximate the infection probability as $v_{j\infty} \approx \frac{1}{n} \sum_{i=1}^{n} E[X_j(t + i\Delta)]$, while the autocorrelation $R_{j\infty}(h)$, which approximates the exact autocorrelation $\rho_j(s, t)$ in (1), is just the autocorrelation of the binary infection sequence $X_j(t + k\Delta)$. Furthermore, using the NIMFA equation in the metastable state $-\delta v_{j\infty} + \beta(1 - v_{j\infty}) \sum_{i=1}^{N} a_{ji} v_{i\infty} = 0$, we can eliminate δ and (22) becomes

$$\beta = -\frac{v_{j\infty}}{\sum_{i=1}^{N} a_{ji} v_{i\infty}} \frac{\ln[R_{j\infty}(h)]}{h}. \qquad (23)$$

Under NIMFA, the curing rate δ can be estimated by (22) without knowing the underlying network. However, to estimate the infection rate β, formula (23) involves the network information. We rewrite (23) as

$$-\frac{v_{j\infty}}{\beta} \frac{\ln[R_{j\infty}(h)]}{h} = \sum_{i=1}^{N} a_{ji} v_{i\infty}$$
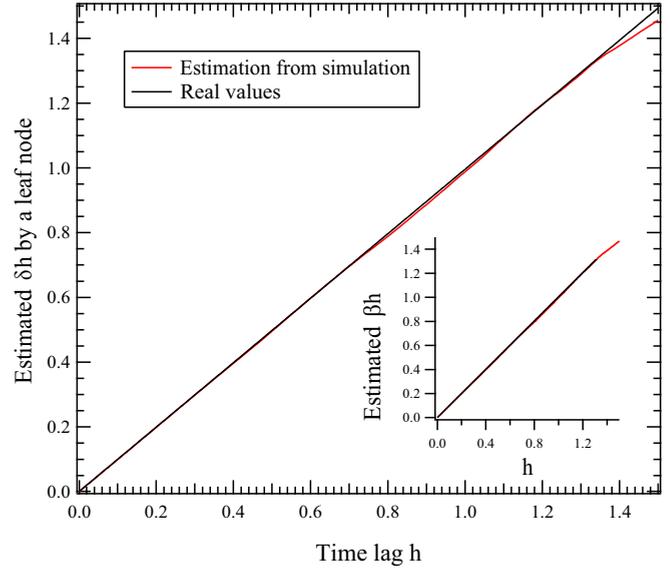
FIG. 8. The estimation of the infection rate β and the curing rate δ using (23) and (22) for the star graph corresponding to Fig. 3. The curves are $-(1 - v_{j\infty}) \ln[R_{j\infty}(h)]$ and $-\frac{v_{j\infty} \ln[R_{j\infty}(h)]}{\sum_{i=1}^{N} a_{ij} v_{i\infty}}$ of a leaf node versus h. Both the estimated β and δ are 1.00, while the real values of rates are both 1.

and sum over all nodes

$$-\sum_{j=1}^{N} \frac{v_{j\infty}}{\beta} \frac{\ln[R_{j\infty}(h)]}{h} = \sum_{j=1}^{N} \sum_{i=1}^{N} a_{ji} v_{i\infty} = \sum_{i=1}^{N} d_i v_{i\infty},$$

where d_i is the degree of node i. After rearrangement of the above equation, we obtain

$$\beta = -\frac{\sum_{j=1}^{N} v_{j\infty}}{\sum_{j=1}^{N} d_j v_{j\infty}} \frac{\ln[R_{j\infty}(h)]}{h}. \qquad (24)$$

Thus, the estimation of the infection rate β requires either the degree of every node d_j for all j as in (24) or the local topology information about node j, i.e., a_{ji} for all i as in (23).

Using the binary infection-state sequence $X_j(t + k\Delta)$ obtained by simulation, we estimate the curing rate δ and the infection rate β by (22) and (23), respectively. In Fig. 8, the value of the estimated rates times the time lag h is plotted for a leaf node of the star graph corresponding to Fig. 3. The slopes of the linear fitting functions (red curves in Fig. 8) are the estimated rates, and both the estimated infection rate β and the curing rate δ are 1.00 while both the real rates equal to 1.

VII. CONCLUSION

In this paper, we study the autocorrelation, the only second-moment property captured by NIMFA, of the SIS process. We obtained the explicit formula of the autocorrelation, i.e., Eq. (8), under NIMFA in the steady state, and the steady-state autocorrelation follows an exponential decay with the time lag. Interestingly, the steady-state autocorrelation is independent of the curing rate δ in regular graphs. Moreover, using the Magnus expansion, we are able to calculate the autocorrelation in the transient state of the SIS process. Our analysis of the transient state not only allows the study of the SIS process above or

below the epidemic threshold but also opens an avenue for the study of the critical behavior [25].

We also evaluated our results by simulation. Although NIMFA assumes that there is no correlation between the infection states of neighbors, i.e., $E[X_i(t)X_j(t)] = E[X_i(t)]E[X_j(t)]$ for $i \neq j$, we show that the NIMFA autocorrelation ($i = j$) is generally accurate by simulation, and the accuracy depends on the accuracy of the NIMFA infection probabilities. If NIMFA can capture the first-order moments, i.e., the infection probability of each node and the prevalence,

under certain SIS parameters and networks, then NIMFA can also be applied to approximate the autocorrelation properties. Finally, we show that our results can be used to estimate the infection and curing rate of the SIS process.

APPENDIX A: AUTOCORRELATION OF $V_{j\infty}$ AND $I_\infty(t)$

When the effective infection rate $\tau > 1/\lambda_1$, the steady infection probability $v_{j\infty}$ can be obtained by solving $dv_j(t)/dt = 0$, i.e.,

$$-\delta v_{j\infty} + (1 - v_{j\infty})\tilde{\beta}_{j\infty} = 0, \tag{A1}$$

where $\tilde{\beta}_{j\infty} = \lim_{t\to\infty} \tilde{\beta}_j(t)$ is time invariant. Thus,

$$\tilde{\beta}_{j\infty} = \frac{\delta v_{j\infty}}{1 - v_{j\infty}} \tag{A2}$$

and the steady infinitesimal generator of node j is

$$Q_{j\infty} \triangleq \lim_{t\to\infty} Q_j(t) = \begin{bmatrix} -\tilde{\beta}_{j\infty} & \tilde{\beta}_{j\infty} \\ \delta & -\delta \end{bmatrix}.$$

In the steady state, the transition probability matrix of $V_{j\infty}(t)$ with time lag h is

$$P_{j\infty}(h) \triangleq \begin{bmatrix} \Pr[V_{j\infty}(t+h)=0|V_{j\infty}(t)=0] & \Pr[V_{j\infty}(t+h)=1|V_{j\infty}(t)=0] \\ \Pr[V_{j\infty}(t+h)=0|V_{j\infty}(t)=1] & \Pr[V_{j\infty}(t+h)=1|V_{j\infty}(t)=1] \end{bmatrix}.$$

By solving the Kolmogorov forward equation $P'_{j\infty}(h) = P_{j\infty}(h)Q_{j\infty}$ given that $P_{j\infty}(0)$ is an identity matrix, we obtain

$$P_{j\infty}(h) = e^{Q_{j\infty}h} = U_j e^{\Lambda_j} U_j^{-1}, \tag{A3}$$

where U_j and Λ_j are the eigenvector matrix and the diagonal eigenvalue matrix of $Q_{j\infty}h$, respectively. The term $Q_{j\infty}h$ can be diagonalized as

$$Q_{j\infty}h = \underbrace{\begin{bmatrix} -\frac{\tilde{\beta}_{j\infty}/\delta}{\sqrt{(\tilde{\beta}_{j\infty}/\delta)^2+1}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{(\tilde{\beta}_{j\infty}/\delta)^2+1}} & \frac{1}{\sqrt{2}} \end{bmatrix}}_{U_j} \underbrace{\begin{bmatrix} -(\tilde{\beta}_{j\infty}+\delta)h & 0 \\ 0 & 0 \end{bmatrix}}_{\Lambda_j} \underbrace{\left(-\sqrt{2}\frac{\sqrt{(\tilde{\beta}_{j\infty}/\delta)^2+1}}{\tilde{\beta}_{j\infty}/\delta+1}\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{(\tilde{\beta}_{j\infty}/\delta)^2+1}} & -\frac{\tilde{\beta}_{j\infty}/\delta}{\sqrt{(\tilde{\beta}_{j\infty}/\delta)^2+1}} \end{bmatrix} \right)}_{U_j^{-1}}.$$

By substituting $e^{\Lambda_j} = \begin{bmatrix} e^{-(\tilde{\beta}_{j\infty}+\delta)h} & 0 \\ 0 & e^0 \end{bmatrix}$ and U_j into (A3), we obtain the steady-state transition probability matrix

$$P_{j\infty}(h) = \begin{bmatrix} \frac{\tilde{\beta}_{j\infty}e^{-(\tilde{\beta}_{j\infty}+\delta)h}+\delta}{\tilde{\beta}_{j\infty}+\delta} & \frac{-\tilde{\beta}_{j\infty}e^{-(\tilde{\beta}_{j\infty}+\delta)h}+\tilde{\beta}_{j\infty}}{\tilde{\beta}_{j\infty}+\delta} \\ \frac{-\delta e^{-(\tilde{\beta}_{j\infty}+\delta)h}+\delta}{\tilde{\beta}_{j\infty}+\delta} & \frac{\delta e^{-(\tilde{\beta}_{j\infty}+\delta)h}+\tilde{\beta}_{j\infty}}{\tilde{\beta}_{j\infty}+\delta} \end{bmatrix}. \tag{A4}$$

From (A4), the joint expectation for $h \geqslant 0$ is

$$E[V_{j\infty}(t)V_{j\infty}(t+h)] = \Pr[V_{j\infty}(t)=1, V_{j\infty}(t+h)=1] = \Pr[V_{j\infty}(t+h)=1|V_{j\infty}(t)=1]\Pr[V_{j\infty}(t)=1]$$

$$= v_{j\infty}\frac{\delta e^{-(\tilde{\beta}_{j\infty}+\delta)h}+\tilde{\beta}_{j\infty}}{\tilde{\beta}_{j\infty}+\delta}. \tag{A5}$$

By substituting $E[V_j(t)V_j(t+h)]$ from (A5) and $\tilde{\beta}_{j\infty}$ from (A2) into (7), we obtain (8).

The autocorrelation of the fraction of infected nodes $I_{j\infty}(t)$ is

$$R_{I\infty}(h) = \frac{E[I_\infty(t)I_\infty(t+h)] - E[I_\infty(t)]E[I_\infty(t+h)]}{\text{Var}[I(t)]} = \frac{\sum_{j=1}^N \left(E[V_{j\infty}(t)V_{j\infty}(t+h)] - v_{j\infty}^2\right)}{\sum_{j=1}^N \left(v_{j\infty} - v_{j\infty}^2\right)}. \tag{A6}$$

From (7), the term $E[V_{j\infty}(t)V_{j\infty}(t+h)] - v_{j\infty}^2 = R_{j\infty}(h)(v_j - v_j^2)$. Thus, we obtain (13).

APPENDIX B: THE HMF AUTOCORRELATION IN THE STEADY STATE

HMF assumes the SIS process is running on an annealed network where nodes with the same degree are statistically equivalent [1] or a time-varying network with infinite rewiring rate [35]. On static networks, NIMFA performs better [15]. The HMF equation is

$$\frac{dw_d(t)}{dt} = -\delta w_d(t) + \beta(1 - w_d(t))d \sum_{k=1}^{N-1} f(k,d)w_k(t),$$

(B1)

where $w_d(t)$ denotes the infection probability of the nodes with degree d, and $f(k,d)$ is the probability that an edge of a node with degree d connects to a node with degree k. The HMF threshold is $\tau_c^{\mathrm{HMF}} = \frac{E[D]}{E[D^2]}$, where D is the degree of a randomly selected node. We assume $\tilde{\beta}_d^{\mathrm{HMF}}(t) = \beta d \sum_{k=1}^{N-1} f(k,d)w_k(t)$. In the steady state $dw_d(t)/dt = 0$ when $\tau > \tau_c^{\mathrm{HMF}}$, the HMF infection probability $w_{d\infty} \triangleq \lim_{t\to\infty} w_d(t)$ follows:

$$-\delta w_{d\infty} + (1 - w_{d\infty})\tilde{\beta}_{d\infty}^{\mathrm{HMF}} = 0,$$

(B2)

where $\tilde{\beta}_{d\infty}^{\mathrm{HMF}} = \lim_{t\to\infty} \tilde{\beta}_d^{\mathrm{HMF}}(t)$. Equation (B2) has a same form with (A1), and the derivation of the HMF autocorrelations is also similar. For example, in the steady state, following the same derivation in Appendix A, the HMF autocorrelation of nodes with degree d is

$$R_{d\infty;\mathrm{HMF}} = e^{-\frac{\delta}{1-w_{d\infty}}h}.$$

The HMF autocorrelation of the fraction of infected nodes is

$$R_{I\infty;\mathrm{HMF}}(h) = \frac{\sum_{d=1}^{N}[\mathrm{Pr}(D=d)]^2 R_{d\infty;\mathrm{HMF}}(w_{d\infty} - w_{d\infty}^2)}{\sum_{d=1}^{N}[\mathrm{Pr}(D=d)]^2(w_{d\infty} - w_{d\infty}^2)}.$$

Here we assume that the HMF fraction of infection nodes is $\sum_{d=1}^{N} \mathrm{Pr}(d)W_d(t)$, where $W_d(t)$ is the infection state of nodes with degree d. The state transition of nodes with the same degree is considered as coupled Markov processes and the infection states of nodes with the same degree are same. In the case of regular graphs, HMF and NIMFA are equivalent and then the two approximate autocorrelations are equal.

APPENDIX C: THE MAGNUS EXPANSION FOR TIME-INHOMOGENEOUS MARKOV PROCESSES

In this section, we shortly introduce the Magnus expansion and then apply the Magnus expansion to NIMFA to derive the necessary results used in analyzing the NIMFA transition matrix.

In a time-inhomogeneous Markov process with D states, the $D \times D$ probability transition matrix $P(s,t)$ from time s to time t follows the Kolmogorov forward equation

$$\frac{dP(s,t)}{dt} = P(s,t)Q(t),$$

(C1)

where $Q(t)$ is the $D \times D$ time-dependent infinitesimal generator. Stroock [36, p. 164] analyzes Eq. (C1) by dividing time into smaller subintervals with length $1/n$. In each subinterval, the infinitesimal generator $Q(t)$ is assumed to be constant and an approximate transition matrix $P^{[n]}(s,t)$ can be obtained. The

transition matrix $P^{[n]}(s,t)$ converges to the unique solution $P(s,t)$ when $n \to \infty$, which follows the time-inhomogeneous Chapman-Kolmogorov equation,

$$P(s,t) = P(s,r)P(r,t),$$

(C2)

for $s \leqslant r \leqslant t$. However, n is always finite for the practical calculation of $P(s,t)$, and Stroock's method do not give a hint on the accuracy of the calculation.

More generally, the linear Eq. (C1) always [17, p. 166] has a unique solution in form $P(s,t) = \exp[\Omega(s,t)]$, where $\Omega(s,t)$ is an $D \times D$ matrix. If the commutative property $Q(t_1)Q(t_2) = Q(t_2)Q(t_1)$ holds for any $t_1,t_2 \in [s,t]$, then $\Omega(s,t) = \int_s^t Q(u)du$ and Eq. (C1) has a closed form solution $P(s,t) = \exp(\int_s^t Q(u)du)$. However, the commutative property does not necessarily hold in most, if not all, time-inhomogeneous Markov processes.

Equation (C1) can be analyzed using the Magnus expansion [16] when $Q(t)$ is not commutative. In a small time interval $t \in [s,s+T]$ such that [17, Theorem 9]

$$\int_s^{s+T} \|Q(t)\|dt < \pi,$$

(C3)

where $\|.\|$ is 2-norm defined for a matrix Q as $\|Q\| \triangleq \max_{\|x\|=1} \|Qx\|$ and for a vector $x = [x_1, \ldots, x_n]^T$ as $\|x\| = \sqrt{\sum_{i=1}^{n} |x_i|^2}$, the matrix $\Omega(s,t)$ can be expanded into a convergent Magnus series $\Omega(t) = \sum_{k=1}^{\infty} \Omega_k(t)$. The convergent condition (C3) is only sufficient but not necessary. The first term of the Magnus expansion of $\Omega(s,t)$ is just the exponent of the solution of (C1) by assuming the commutative property of $Q(t)$, i.e.,

$$\Omega_1(s,t) = \int_s^t Q(u)du.$$

(C4)

The second term of the Magnus series is

$$\Omega_2(s,t) = \frac{1}{2} \int_s^t du_1 \int_s^{u_1} du_2 [Q(u_1), Q(u_2)],$$

where $[A,B] \triangleq AB - BA$ is the matrix commutator, and the third term is

$$\Omega_3(s,t) = \frac{1}{6} \int_s^t du_1 \int_s^{u_1} du_2 \int_s^{u_2} du_3([Q(u_1),$$
$$[Q(u_2),Q(u_3)]] + [Q(u_3),[Q(u_2),Q(u_1)]]).$$

The calculation of further terms can be found in Ref. [17], which is not involved in this paper.

The transition matrix $P(s,t)$ is time symmetric in the sense that $P(s,t)P(t,s) = I$ and then $\Omega(s,t) = -\Omega(t,s)$. In the time interval $[s,t]$,

$$\Omega\left(\frac{s+t}{2} - h, \frac{s+t}{2} + h\right) = -\Omega\left(\frac{s+t}{2} + h, \frac{s+t}{2} - h\right)$$

for $h \leqslant (t-s)/2$ and thus the odd function $\Omega_k(\frac{s+t}{2} - h, \frac{s+t}{2} + h)$ only contains odd powers of h in its Taylor expansion [17, p. 165]. Thus, $\Omega_{2k}(s,t) = O[(t-s)^{2k+1}]$ and $\Omega_{2k+1}(s,t) = O[(t-s)^{2k+3}]$ in the time interval $[s,t]$. Correspondingly, $\sum_{k=1}^{2i+1} \Omega_k(s,t) = \Omega(s,t) + O[(t-s)^{2i+3}]$ and $\sum_{k=1}^{2i} \Omega_k(s,t) = \Omega(s,t) + O[(t-s)^{2i+3}]$. The sums of the

first $2i$ and $2i+1$ terms of the Magnus series of $\Omega(s,t)$ achieve a same order of accuracy with respect to the time length $t-s$. Moreover, the power series of the matrix exponential indicates that $\exp\{\Omega(s,t) + O[(t-s)^k]\} = \exp[\Omega(s,t)] + O[(t-s)^k]$. Specifically, we have $P(s,t) = \exp[\Omega_1(s,t)] + O[(t-s)^3]$ by only keeping the first term in the Magnus expansion. Using the Taylor expansion of $\Omega_1(s,t)$, we may find that Stroock's method only achieves a second-order accuracy by only preserving the first term of the Taylor series of $Q(t)$.

In NIMFA, given the infection probabilities $v_i(t)$ for $i = 1,\ldots,N$, the infinitesimal generator $Q_j(t)$ of a node j defined by (6) is a determined function of time. Thus, we can apply the Magnus expansion to the Markov process (5) and assume the transition matrix $P_j(s,t) = \exp[\Omega(s,t;j)]$. First, we derive the length of the convergent time interval of the

Magnus expansion by condition (C3). We may verify that the 2-norm of the matrix $Q_j(t)$ is $\|Q_j(t)\| = \sqrt{\tilde{\beta}_j^2(t) + \delta^2}$. For $t > 0$, $\tilde{\beta}_j(t) = \beta \sum_{i=1}^{N} a_{ji} v_i(t) < \beta d_j$, where d_j is the degree of node j, and, consequently, $\|Q_j(t)\| < \sqrt{\beta^2 d_j^2 + \delta^2}$. Thus,

$$\int_s^{s+T} \|Q_j(t)\| dt < \sqrt{\beta^2 d_j^2 + \delta^2}\, T.$$

Let $\sqrt{\beta^2 d_j^2 + \delta^2}\, T = \pi$, and thus

$$T = \frac{\pi}{\sqrt{\beta^2 d_j^2 + \delta^2}}. \tag{C5}$$

The Magnus expansion of $\Omega(s,t;j)$ always converges if $t - s \leqslant T$ from Eq. (C3).

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

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

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

[4] P. Van Mieghem, arXiv:1310.3980 [math.PR] (2013).

[5] P. Donnelly, Math. Biosci. **117**, 49 (1993).

[6] R. M. Anderson, B. T. Grenfell, and R. M. May, J. Hygiene **93**, 587 (1984).

[7] P. L. Simon, M. Taylor, and I. Z. Kiss, J. Math. Biol. **62**, 479 (2011).

[8] J. Meier, X. Zhou, A. Hillebrand, P. Tewarie, C. Stam, and P. Van Mieghem, NeuroImage **152**, 639 (2017).

[9] P. Van Mieghem, *Performance Analysis of Complex Networks and Systems* (Cambridge University Press, Cambridge, 2014).

[10] P. Van Mieghem, Delft University of Technology, Report 20170405 (2017), https://www.nas.ewi.tudelft.nl/people/Piet/TUDelftReports.html.

[11] NIMFA is sometimes called quenched mean-field approximation (QMF) or individual-based mean-field approach (IBMF) as in Ref. [1].

[12] P. Van Mieghem, Delft University of Technology, Report 20180312 (2018), https://www.nas.ewi.tudelft.nl/people/Piet/TUDelftReports.html.

[13] A. Khanafer, T. Basar, and B. Gharesifard, in *American Control Conference (ACC'14)* (IEEE, Los Alamitos, CA, 2014), pp. 3579–3584.

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

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

[16] W. Magnus, Commun. Pure Appl. Math. **7**, 649 (1954).

[17] S. Blanes, F. Casas, J. Oteo, and J. Ros, Phys. Rep. **470**, 151 (2009).

[18] Q. Liu and P. Van Mieghem, Physica A **471**, 325 (2017).

[19] R. Van de Bovenkamp, Epidemic processes on complex networks: Modelling, simulation and algorithms, Ph.D. thesis, TU Delft, Delft University of Technology, 2015.

[20] W. Cota and S. C. Ferreira, Comput. Phys. Commun. **219**, 303 (2017).

[21] P. Van Mieghem and R. van de Bovenkamp, Phys. Rev. E **91**, 032812 (2015).

[22] J. P. Gleeson, S. Melnik, J. A. Ward, M. A. Porter, and P. J. Mucha, Phys. Rev. E **85**, 026106 (2012).

[23] K. Devriendt and P. Van Mieghem, Phys. Rev. E **96**, 052314 (2017).

[24] Q. Liu and P. Van Mieghem, in *Complex Networks & Their Applications V*, edited by H. Cherifi, S. Gaito, W. Quattrociocchi, and A. Sala (Springer International Publishing, Cham, 2017), pp. 511–521.

[25] For example, evidence has shown that there exists an extended critical region just above the NIMFA threshold $\tau_c^{(1)}$ in some networks: the Griffiths phase [26-28] related to the epidemic localization [29-32] evaluated by the behavior around the NIMFA threshold [33]. With the extended critical region, the effective infection rate τ needs not to be fine-tuned to the exact threshold to let the process be critical [34]. In the critical region, the NIMFA steady-state prevalence $y_\infty^{(1)} \triangleq \lim_{t\to\infty} y^{(1)}(t)$ is very small (converging to zero with the increase of network N), and it is similar to that below the threshold $\tau_c^{(1)}$. However, the prevalence is different in the transient state: The prevalence follows a power-law decay in the critical region [27] while it follows an exponential decay below the threshold. The critical autocorrelation properties in the Griffiths phase can be studied with our analysis of the transient state.

[26] H. K. Lee, P.-S. Shim, and J. D. Noh, Phys. Rev. E **87**, 062812 (2013).

[27] W. Cota, S. C. Ferreira, and G. Ódor, Phys. Rev. E **93**, 032322 (2016).

[28] M. A. Muñoz, R. Juhász, C. Castellano, and G. Ódor, Phys. Rev. Lett. **105**, 128701 (2010).

[29] A. V. Goltsev, S. N. Dorogovtsev, J. G. Oliveira, and J. F. F. Mendes, Phys. Rev. Lett. **109**, 128702 (2012).

[30] R. Pastor-Satorras and C. Castellano, J. Stat. Phys. (2018), doi:10.1007/s10955-018-1970-8.

[31] F. D. Sahneh, A. Vajdi, and C. Scoglio, in *American Control Conference (ACC'16)* (IEEE, Los Alamitos, CA, 2016), pp. 7346–7351.

[32] R. S. Ferreira, R. A. da Costa, S. N. Dorogovtsev, and J. F. F. Mendes, Phys. Rev. E **94**, 062305 (2016).

[33] P. Van Mieghem, Europhys. Lett. **97**, 48004 (2012).

[34] P. Moretti and M. A. Muñoz, Nat. Commun. **4**, 2521 (2013).

[35] G. St-Onge, J.-G. Young, E. Laurence, C. Murphy, and L. J. Dubé, Phys. Rev. E **97**, 022305 (2018).

[36] D. W. Stroock, *An Introduction to Markov Processes*, Vol. 230 (Springer Science & Business Media, Berlin, 2013).