Linking population-level models with growing networks: A class of epidemic models

Romulus Breban, Raffaele Vardavas, and Sally Blower
Department of Biomathematics, University of California, Los Angeles, California 90095-1555, USA
(Received 2 June 2005; published 11 October 2005)

We introduce a class of growing network models that are directly applicable to epidemiology. We show how to construct a growing network model (individual-level model) that generates the same epidemic-level outcomes as a population-level ordinary differential equation (ODE) model. For concreteness, we analyze the susceptible-infected (SI) ODE model of disease invasion. First, we give an illustrative example of a growing network whose population-level variables are compatible with those of this ODE model. Second, we demonstrate that a growing network model can be found that is equivalent to the Crump-Mode-Jagers (CMJ) continuous-time branching process of the SI ODE model of disease invasion. We discuss the computational advantages that our growing network model has over the CMJ branching process.

DOI: 10.1103/PhysRevE.72.046110 PACS number(s): 89.75.Hc, 87.19.Xx, 87.23.Ge, 05.10.−a

I. INTRODUCTION

Networks arise in many natural and artificial systems. In the social and biological sciences, networks are used to model the contacts among individuals. It has long been recognized that networks can naturally describe the contact dynamics of a spreading disease; for example, SARS, HIV, and others. Previously, branching processes have been used to specify the transmission network at the beginning of a disease outbreak (i.e., disease invasion) [1–8]. The individuals are represented by nodes which are directionally linked, indicating who infected whom. The most accepted branching process applied to epidemiology is known as the Crump-Mode-Jagers (CMJ) process [2–8]. This model has been used in the modeling of the HIV and malaria epidemics [6]. In the past decade, a new field of network models has been developed for modeling the spread of infectious diseases. In these models, a given contact network is coupled with a probabilistic transmission model: every infected node (individual) has a given probability for infecting any of its susceptible (i.e., noninfected) neighbors in the contact network [9–23]. The disease invades the contact network through percolation, and finding the percolation thresholds becomes of capital importance in predicting the growth of an epidemic.

Many epidemiological models [24] are based on ordinary differential equations (ODEs). ODE epidemic models successfully describe the observed population-level data (such as the number of new reported cases per unit time; see e.g., [24–26]); however, these aggregate data are far from sufficient for validating a network model. The invasion of an epidemic is now modeled as a threshold phenomenon (i.e., typically a transcritical bifurcation) characterized by a standard parameter called the basic reproduction ratio (or number) R_0. On general grounds, R_0 is biologically defined at disease invasion[1] as follows. R_0 represents the number of secondary cases caused by an infected individual during his/her entire infectious period [27–31]. It is understood that if $R_0>1$ then a disease outbreak settles to an endemic level, or if $R_0<1$, the disease outbreak goes extinct. R_0 is a critical parameter which is used not only to assess epidemic severity, but also to design control strategies.[2] In practice, R_0 can be estimated through contact tracing.[3] In theory, several methods have been used for calculating R_0 for ODE models: (i) stability analysis (e.g., [25]), (ii) the "fraction of the host population that is susceptible at equilibrium" [24], and (iii) the next generation analysis [24,27–31]. However, these methods make further individual-level assumptions without explicitly writing a full individual-level model (ILM). We argue that, in order to associate an R_0 to an ODE model, an ILM which is compatible to the ODE model must be developed; only then can the R_0 of the ILM be unambiguously calculated. In this work we develop a mathematical setup where we can directly apply the definition of R_0. Then, we verify how R_0 satisfies the presumed threshold property.

Population-level epidemic data can be described through network models; the expected number of network nodes in given disease states specify the epidemic state at the population level. Measurements on epidemic networks (e.g., [32,33]), and increased computational power for storage and manipulation have increased the opportunity for developing new realistic models of the network evolution. Evidently, these network models will have to be compatible with their predecessors based on ODEs. In principle, there are two major approaches to this problem. On one hand, a network model could be built from individual-level assumptions such that the corresponding ODE model results as a by-product of statistics on the network (e.g., CMJ constructs [2–8]); in this

[1]The case of disease invasion assumes that an infected individual encounters only susceptibles, and that there is no depletion of susceptibles.

[2]The larger is the value of R_0 for an epidemic, the harder it is to eradicate that epidemic. Also, an R_0 formula provides control parameters and suggests public health policies to decrease R_0 below the threshold.

[3]Say, for example, that the end of the infectious period of an individual is marked by hospital self-check-in. The individual coming into the hospital may be interviewed for his/her recent contacts. Then, the number of the contacts who are infected with the same disease can be counted. By averaging over many individuals that check themselves into the hospital, an R_0 estimate can be obtained.

case, calculating the expected number of nodes in given disease states versus time yields a solution of the ODE model of interest. On the other hand, a network model could be constrained to be compatible with a given ODE model. Here we follow the second path and build a class of network models that are constrained to be compatible with a chosen ODE model. Our construct is based on the kinetic Monte Carlo (KMC) algorithm [34,35] and sets of rules for growing a disease transmission network; for literature on growing network topology see [36–44]. We call our new constructs *KMC growing networks*. As an example of an ODE which can be assigned a class of KMC growing networks, we discuss the *susceptible-infected* (SI) ODE model of disease invasion. Every network model in its class has a choice of two rules for growing the transmission network: an *infection rule* and a *removal rule*. We first present our network class in a systematic fashion discussing the KMC algorithm. Then we present analytical and numerical results on a particular example. Finally, we address the very important question of compatibility between the CMJ and the KMC network dynamics. We make the following two points. First, given an ODE epidemic model, there is an infinite class of KMC growing networks that are all compatible with the ODE; the CMJ process, if explicitly constructed, is unique. Second, for our SI ODE model example, we show that a slightly modified KMC growing network scheme yields identical dynamics to that of the corresponding CMJ process. The overall purpose of this work is to define KMC growing networks and evaluate their versatility and modeling potential. In doing so we briefly analyze the concept of R_0; a more comprehensive analysis of R_0 will be given elsewhere [45].

II. CLASS OF INDIVIDUAL-LEVEL MODELS

We consider the SI model at disease invasion as our paradigm ODE model. The classic SI model is given as follows:

$$dS/dt = -\beta SI/N,$$

$$dI/dt = \beta SI/N - \mu I, \tag{1}$$

where $S(t)$, $I(t)$, and $N(t)=S(t)+I(t)$ are the susceptible, infected, and total populations at time t, respectively. β denotes the infectiousness of the disease, and μ denotes the per capita removal rate due to the disease. It is very important to note that β and μ are in general population-level averages, and for a heterogeneous population they may provide a poor description for the rates of typical individual-level processes. At disease invasion, we assume that depletion of susceptibles is negligible (i.e., $S/N \approx 1$) and thus obtain

$$dI/dt = \beta I - \mu I, \tag{2}$$

which is the SI ODE model at disease invasion.

A compartmental ODE model has a naturally assigned continuous-time finite Markov chain [46,47]. The expectation values of the populations in the Markov chain compartments (i.e., the mean field) in the limit of large populations yield the ODE solution.[4] The Markov chain corresponding to the SI ODE model of disease invasion is a simple birth-death process [47]. The KMC algorithm for integrating this Markov chain consists of the repetition of three steps that update the number of infected individuals and time.

(1) Randomly select a process which is either the infection of an individual or the removal of an already infected individual. A new infection occurs with rate βI and a removal occurs with rate μI. The total process rate is $\mathcal{R}=(\beta+\mu)I$. A process is selected with probability given by its rate normalized by $\mathcal{R}$. Therefore, an infection occurs with probability $\beta/(\beta+\mu)$ and a removal with probability $\mu/(\beta+\mu)$.

(2) Update the value of I as follows. If a new infection occurs then $I \to I+1$, otherwise a removal occurs and $I \to I-1$.

(3) Update time. Since Markov processes are memoryless, the interevent time is exponentially distributed with average $\mathcal{R}^{-1}$. Thus, increment time by $\delta t = -(\ln U)/\mathcal{R}$ where U is a uniform random variable in [0, 1].

In addition to the described KMC algorithm, we track individuals in the compartments. Thus, we construct an ILM by implementing a growing network scheme of who infected whom. Individuals are represented by nodes in the network, and two individuals A and B are connected by a directed link from B to A if B has infected A. Under these circumstances, step 2 of the above algorithm is modified as follows.

(2′) If a new infection occurs, a new node A with no prior links to the growing network is added in the following way. Choose a network node B in the infected compartment according to a certain rule which we call *the infection rule*. Add a directed link from B to A meaning that B has infected A; see Fig. 1. By default, A belongs to the infected compartment until it is removed. If a removal occurs, choose a network node C in the infected compartment according to a certain rule which we call *the removal rule*, and remove that node from the infected compartment. The node C does remain connected to the network but it is not available to be assigned new connections; see Fig. 1. The count of nodes in the infected compartment yields the number of infected individuals I. As a consequence of the above procedure, we have that $I \to I+1$ if an infection occurs, and that $I \to I-1$ if a removal occurs.

This construct provides a class of ILMs fully compatible with the SI ODE model of disease invasion since the tracking of individuals in step 2′ does not alter the dynamics of I in step (2) of the KMC algorithm. The infection and the removal rules remain to be specified, and each distinct set of rules yields a distinct ILM in this class. This is an explicit construct of a class of ILMs which are compatible with the

[4] In general, compatibility between an ODE model and its corresponding continuous-time Markov chain requires the limit of large population. However, this is not the case for the SI ODE model of disease invasion and its corresponding Markov chain. This relates to the fact that the SI ODE model of disease invasion is linear. It can be verified directly through moment closure approximations of the Markov chain corresponding to Eq. (2) that the compatibility of the SI ODE model of disease invasion and its corresponding Markov chain holds for all population sizes.

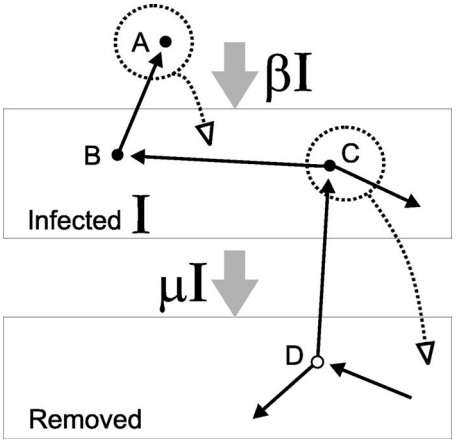

FIG. 1. Structure of the KMC model with tracking of individuals through growing networks. Individuals are the nodes of the network and two individuals A and B are connected by a directed link from B to A if B has infected A. A is a newly infected individual added to the transmission growing network. By the *infection rule*, B is chosen to be the one who has infected A. If a removal occurs, an individual C is chosen in the infected compartment according to the *removal rule*, and then C is removed from the infected compartment. The node C does remain connected in the network but it is not available to be assigned new connections; i.e., as node D marked by an open circle. The rates of infection and removal are dictated by the KMC algorithm applied to the SI ODE model of disease invasion.

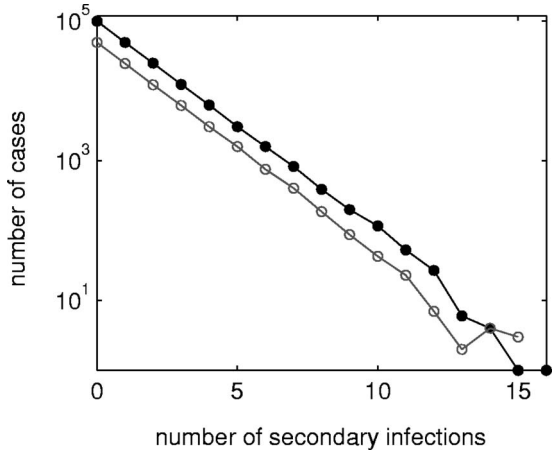

FIG. 2. The unnormalized degree distributions of removed nodes (black disks) and infectious nodes $I_Q(500\,000)$ (open circles) resulting from the ILM with purely stochastic infection and removal rules (see Sec. III) which was run for a total of $n=500\,000$ events. The parameters of the computation are $\beta=0.015$ and $\mu=0.01$. The featured variations are approximately exponential; in semilogarithmic scale, the graphs appear to be straight lines with the slopes $-0.296\ldots\approx-\log_{10}(2)$ (black disks) and $-0.308\ldots\approx-\log_{10}(2)$ (open circles).

same ODE model. Our ILM class belongs to the very broad class of growing networks since individuals are added to a network of who infected whom based on global or local network rules. However, given that a newly infected individual is connected by a single directed link to the existent network, our network contains only trees and no loops, and it has the network topology of branching processes. The average number of secondary infections of removed individuals R_0 is computed as the average number of outgoing links of a removed node (i.e., a node that no longer accepts new links). The average number of secondary infections of actively infectious individuals $Q_0(t)$ is computed as the average number of outgoing links of a node in the infected compartment at time t. On very general grounds, $Q_0(t)$ and $R_0(t)$ satisfy a constraint that we discuss in Appendix A.

The SI ODE model of disease invasion that we analyze in this work is simple, and yet illustrative. Our methods of analysis are not restricted to this ODE model. The generalization of our individual-level construct to an arbitrary compartmental ODE model is conceptually straightforward as every such model has naturally assigned a Markov chain which can be integrated by KMC methods. The corresponding KMC algorithm is then modified to fully track individuals by growing a transmission network which is implemented by individual-level rules. We thus believe that the ideas we have presented here apply to a very large class of compartmental ODE models.

III. A KMC GROWING NETWORK EXAMPLE

We now present an example ILM with the following infection and removal rules. *Infection rule:* An individual joining the infectious pool is infected by an infectious individual who is uniformly randomly selected. *Removal rule:* When a removal occurs, a uniformly randomly selected individual leaves the infectious pool. Our interest is in the steady state degree distribution of the infectious nodes. We denote the expected number of nodes in the infectious pool having Q outgoing connections after n processes by $I_Q(n)$. The open circles in Fig. 2 show numerical results for $I_Q(n)$ versus Q at large n (i.e., $n=500\,000$). These suggest that $I_Q(n)$ at large n varies exponentially with Q, and that the rate of the exponential variation is approximately $-\log_{10}(2)$. This observation can be confirmed analytically as follows.

The rate equations for the expected number of nodes with $Q\geqslant 0$ connections are

$$I_0(n+1)=I_0(n)+\frac{\beta}{\beta+\mu}-\frac{\beta}{\beta+\mu}\frac{I_0(n)}{I(n)}-\frac{\mu}{\beta+\mu}\frac{I_0(n)}{I(n)}, \tag{3}$$

$$I_Q(n+1)=I_Q(n)+\frac{\beta}{\beta+\mu}\frac{I_{Q-1}(n)}{I(n)}-\frac{\beta}{\beta+\mu}\frac{I_Q(n)}{I(n)}-\frac{\mu}{\beta+\mu}\frac{I_Q(n)}{I(n)}, \tag{4}$$

where $I(n)$ is the expected total number of nodes in the infected pool after n processes. $I(n)$ is given by

$$I(n)=I(0)+n\frac{\beta-\mu}{\beta+\mu}. \tag{5}$$

$I_0(n+1)$ is the sum of $I_0(n)$ with the following three terms. The first term is the probability that event $(n+1)$ is an infection, and that the newly infected individual joins the $I_0(n)$

category. The second term represents the probability that event $(n+1)$ is an infection and that the infection was caused by an individual in the $I_0(n)$ category who will be transfered to the $I_1(n)$ category. The last term represents the probability that event $(n+1)$ is a removal from the $I_0(n)$ category. Eq. (4) contains similar terms except for the first one which represents the inflow in the $I_Q(n)$ category due to the fact that event $(n+1)$ is an infection due to an individual in the $I_{Q-1}(n)$ category.

Denote by $\rho_Q(n)$ the probability that a randomly chosen node in the infected pool after n processes has Q links. Using $I(n)\rho_Q(n)=I_Q(n)$ and Eqs. (3) and (4), we obtain

$$\rho_0(n+1) = \rho_0(n)\frac{I(n)}{I(n+1)} + \frac{1}{I(n+1)}\left(\frac{\beta}{\beta+\mu} - \rho_0(n)\right), \tag{6}$$

$$\rho_Q(n+1) = \rho_Q(n)\frac{I(n)}{I(n+1)} + \frac{1}{I(n+1)} \times\left(\rho_{Q-1}(n)\frac{\beta}{\beta+\mu} - \rho_Q(n)\right). \tag{7}$$

These equations give an infinite-dimensional iterated map for the evolution of ρ_Q. Since $1/I(n)\to 0$ and $I(n)/I(n+1)\to 1$ as $n\to\infty$, $\rho_Q(n)$ converges to a definite value in $[0, 1]$ as $n\to\infty$. Eq. (6) can be solved for $\rho_0(n+1)$ as

$$\rho_0(n+1) = \rho_0(0)\prod_{i=0}^{n} a_i + \left(\sum_{i=0}^{n-1}\prod_{j=i+1}^{n} a_j + b_n\right), \tag{8}$$

where

$$a_i \equiv \frac{I(i)}{I(i+1)} - \frac{1}{I(i+1)}, \tag{9}$$

$$b_i \equiv \frac{\beta}{\beta+\mu}\frac{1}{I(i+1)}. \tag{10}$$

The term in parentheses in the right-hand side of (8) can be written using the Euler Γ function as

$$\left(\sum_{i=0}^{n-1}\prod_{j=i+1}^{n} a_j + b_n\right) = \frac{\beta}{\beta+\mu}\left\{\frac{1}{I_0+\alpha(n+1)} + \frac{\Gamma(1+n+(I_0-1)/\alpha)}{(1+\alpha)\Gamma(2+n+I_0/\alpha)}\left(\frac{\Gamma(1+n+I_0/\alpha)}{\Gamma(n+(I_0-1)/\alpha)} - \frac{\Gamma(1+I_0/\alpha)}{\Gamma((I_0-1)/\alpha)}\right)\right\}, \tag{11}$$

where $I_0\equiv I(0)$ and $\alpha\equiv(\beta-\mu)/(\beta+\mu)$. Using that $\Gamma(x)\sim\sqrt{2\pi}x^{x-1/2}\exp(x)$ as $x\to\infty$, we find that the limit of Eq. (11) as $n\to\infty$ is $1/2$. From Eq. (5) we obtain that, for large i, $\ln(a_i)\sim -1/i$, and thus

$$\prod_{i=0}^{\infty} a_i = \exp\left(\sum_{i=0}^{\infty}\ln a_i\right) \sim \exp\left(-\sum_{i=0}^{\infty} 1/i\right) \to 0. \tag{12}$$

Therefore, from Eq. (8) and using the above results, we obtain

$$\rho_0(\infty) = 1/2. \tag{13}$$

Using $\rho_{Q-1}(\infty)$ instead of $\rho_{Q-1}(n)$ in the iterated equation of $\rho_Q(n)$ [Eq. (7)] we obtain the following recursion relation for all $\rho_Q(\infty)$:

$$\rho_Q(\infty) = (1/2)\rho_{Q-1}(\infty), \tag{14}$$

and thus

$$\rho_Q(\infty) = \frac{1}{2^{Q+1}}. \tag{15}$$

This analytic result is in agreement with the numerics presented in Fig. 2. The slope of the $I_Q(\infty)$ graph in log-linear scale is $-\log_{10}(2)$. Using the expression for $\rho_Q(\infty)$, it becomes easy to calculate the basic reproduction ratio R_0. The average number of outgoing links as $t\to\infty$ is $Q_0=\sum_{Q=0}^{\infty}Q\rho_Q(\infty)=1$ and then, using Eq. (A3), $R_0=1$. Furthermore, since the removal rule is uniformly random, the degree distribution of the removed nodes must be the same as that of the infectious nodes $\rho_Q(\infty)$. An alternate explanation for the result that $R_0=1$ is given in Appendix C. This R_0 result is quite intriguing when analyzed from the traditional epidemiological perspective since we can have epidemic growth (i.e., $\beta>\mu$), but nevertheless $R_0=1$ independently of β and μ. Thus, in this case, R_0 does not signal epidemic growth as anticipated in the Introduction.

IV. KMC GROWING NETWORKS VERSUS THE CMJ PROCESS

A. The KMC network example versus the CMJ process

The CMJ process compatible to the SI ODE model of disease invasion (2) is a continuous-time branching process that has the following stochastic protocol for each individual (i.e., node). Each infectious individual is randomly assigned a time interval for being infectious from an exponential dis-

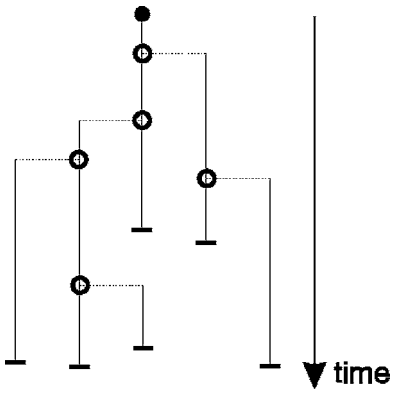

FIG. 3. Schematics of the CMJ branching process. Two generations of a CMJ process are shown. The root is represented by a black disk. Every vertical line segment represents the time interval of being infectious of a certain individual (i.e., node); these time intervals follow a negative exponential distribution with the an average of $1/\mu$. The removal events are marked with dash. The time that infections occurred are marked with open circles. The time intervals between infections for a given individual follow a negative exponential distribution with the an average of $1/\beta$.

tribution with an average of $1/\mu$. The individual has infectious contacts with others and s/he creates new nodes at time intervals which are exponentially distributed with an average of $1/\beta$. Using a well known theorem [48], the CMJ process can also be implemented as follows. Given that an individual stays infectious for a time interval τ, choose the number of infectious contacts from a Poisson distribution with an average of $\beta\tau$ and then choose the times of the contacts uniformly random in the interval $(0, \tau)$. As a result, the CMJ process has $R_0^{\mathrm{CMJ}}=\beta/\mu$ [6]. The probability of extinction of the CMJ process is μ/β when $\beta>\mu$, and 1 when $\beta<\mu$ [6]. For a schematics of the CMJ process, see Fig. 3.

It is important to note that in both the CMJ algorithm and the KMC algorithm presented in Sec. III the individual-level probabilistic processes satisfy the Markov property. Let us consider one individual in the CMJ process. During any infinitesimal time step δt the individual has a probability $\beta\delta t$ to infect and a probability $\mu\delta t$ to be removed from the infectious pool. These probabilities are independent of time and of the past history of the individual. This applies to every individual in the infectious pool and it represents a Markov property that allows for individuals to undergo realizations of the same stochastic scenario (for more involved arguments see [3]). Thus, individuals are statistically identical. However, at the collective level, the CMJ process distinguishes between individuals according to their histories. As a consequence, an individual that has long been infectious has a greater probability of being removed than a recently infected individual.

Let us consider the KMC growing network algorithm presented in Sec. III. During any infinitesimal time step δt the there is a probability $\beta I\delta t$ that an infection process occurs and a probability $\mu I\delta t$ that a removal process occurs. Assuming that the individuals are indistinguishable in the infectious pool, these probabilities are uniformly distributed amongst the individuals such that they become $\beta\delta t$ and $\mu\delta t$ per individual, respectively. These individual-level probabilities are time independent and have the Markov property. Therefore, as with the CMJ process, individuals undergo realizations of the same stochastic scenario and thus they are statistically identical.

The average number of seconday infections over both the infectious and the removed pools is 1 [see Eq. (A1)]. However, in the KMC algorithm presented in Sec. III the process of choosing an individual in the infectious pool is a source of stochasticity that is not present in the CMJ process. Thus, the distributions of the secondary infections of the two processes have different variances. The two processes distribute individual-level fluctuations differently between the infectious pool and the removed pool. In particular, they produce different individual-level averages over the infectious pool only.

B. Construction of a KMC growing network compatible to the CMJ process

The KMC growing networks provide many ILMs which are compatible with the SI ODE model of disease invasion (2). Since the dynamics of I versus the count of individual-level processes is a biased random walk, applying the gambler's ruin theorem [47], we obtain that the extinction probability of all the KMC processes is just the same as the extinction probability of the CMJ process. While $R_0^{\mathrm{CMJ}}=\beta/\mu$, a variety of R_0 values are possible for KMC growing networks (see Sec. III and Appendix A). The class of KMC growing networks is very broad and it is even possible to construct a set of rules such that the resulting KMC growing network is equivalent to the CMJ process, as we demonstrate below.

Our KMC growing network is constructed based on the following infection rule: an individual joining the infectious pool is infected by an individual who is uniform randomly selected from the infectious pool. The removal is done as follows. Every infected individual is assigned a time interval of being infectious which is exponentially distributed with an average of $1/\mu$. At the end of this time interval, the individual is removed from the infectious pool. Figure 4(a) shows numerical results from implementing this growing network algorithm. Each plot, shows the average R_0 values versus the date at which individuals joined the infectious pool. We show three curves corresponding to three different times of running the KMC growing network. As expected, as the running time of the process increases, R_0 converges to the value β/μ. In Fig. 4(b) we compare three graphs that were generated for the same running time. The first graph, plotted with dots, is created with the KMC growing network. The second graph, plotted with open circles, is created with the CMJ process. The third graph, shown as a solid line, is obtained from the following theoretical consideration. Given that individuals remain infectious for a time τ that is exponentially distributed with an average of $1/\mu$ and given that the epidemic process has run until time $T>t$, the average number of secondary infections of an individual that has been infected at time t is given by

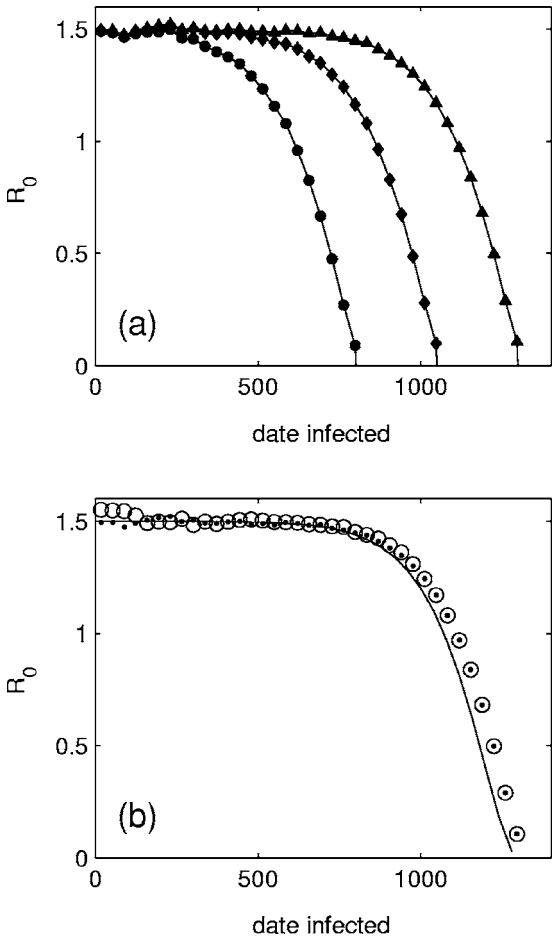

FIG. 4. (a) The distribution of the average R_0 versus the date of infection for three successive running times of the KMC growing network with durations $T=800$ (solid disks), 1050 (diamonds), and 1300 (triangles). The parameters of the computation are $\beta=0.015$ and $\mu=0.01$. Each curve shows the average results over 15 000 runs. The position of each symbol in a curve represents an average over a binning interval of approximate size 35.5. Note that with an increasing running time of the KMC growing network, R_0 approaches β/μ. (b) Comparison of the average R_0 versus the time of infection. The dots represent the results of the KMC growing network, the open circles are the results of the CMJ process, and the line represents a theoretical result.

$$R_0(t;T) = \int_0^{T-t} (\beta\tau)\mu e^{-\mu\tau} d\tau = \frac{\beta}{\mu}\{1 + e^{\mu(t-T)}[\mu(t-T)-1]\}. \tag{16}$$

As seen in Fig. 4(b), all three histograms overlap, numerically indicating that the KMC growing network and the CMJ process perform similarly, and that they are consistent with the theoretical considerations.

In fact, this KMC growing network and the CMJ branching process are indeed equivalent as suggested by the numerics in Fig. 4. The removal of an individual from the infectious pool as implemented in the KMC process is identical to that of the CMJ process. Although the KMC and the CMJ processes are inherently different in treating the infection, their expected outcomes are identical. In both the KMC and the CMJ algorithms the individual-level protocols are independent and statistically identical such that the transmissibility of the disease per individual is β. Generating the infection time intervals can be done either explicitly as in the case of the KMC algorithm or implicitly as in the case of CMJ algorithm [48]. However, the KMC algorithm has practical advantages over the CMJ algorithm. All dynamics in epidemiology evolve over time and predictions need to be made at a given moment of time. The CMJ algorithm is evolved node by node, generation by generation, and individuals from the same generation may occur at very different moments of time; see Fig. 3. Consequently, it is not possible to guarantee that, after evolving a finite number of generations, all the branches of the CMJ process have been fully evolved up to a given moment of time. In contrast, the KMC algorithm is evolved in time, process by process, and this difficulty does not occur.

V. DISCUSSION AND CONCLUSIONS

We find that a broad range of R_0 values are compatible with a given ODE model. Even though an ODE model may successfully fit population-level data, obtaining R_0 from the ODE model is not possible. Since populations are typically heterogeneous, we expect that the R_0^{CMJ} value does not typically apply; this hypothesis could be verified as follows. During an outbreak epidemiologists determine who has infected whom from tracing the contacts of infected individuals and directly estimate R_0. This estimate could then be compared with R_0^{CMJ}. We believe that the two different ways of obtaining an R_0 would disagree even though they would be compatible to the same population-level model. We thus conclude that R_0^{CMJ} associated with the real-life epidemics of infectious diseases such as HIV, SARS, TB, and smallpox is not a suitable parameter for the comparison of their relative severities. In order to justify the use of R_0^{CMJ} for real-life epidemiology, one would have to verify from field data the degree of homogeneity of the populations under epidemiological observation. In this work we show that an infinity of R_0 values may be compatible with an ODE model (see Secs. II, A and [45]). Additional individual-level assumptions are needed to construct an ILM which is compatible with the ODE model, and such assumptions may be supported by epidemic network data. Then, R_0 can directly be calculated from the ILM.

In this work, we have presented a class of individual-level models based on the KMC algorithm and rules for growing a disease transmission network. Our models have the desirable feature that are compatible with a given ODE system. We have presented the situation where the ODE model of interest is the SI ODE model at disease invasion, and we have discussed in detail a KMC growing network example. We also analyzed another KMC network model which we showed to be equivalent to the CMJ process corresponding to the same ODE model. Given that an ODE population model is naturally assigned a Markov chain, our methods would easily apply to a large class of ODE models. We conclude that KMC growing networks are powerful tools of great modeling potential for epidemiology.

ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support from NIH (Grant No. RO1 AI041935). We also thank Erin Bodine, Robert Smith, and David Wilson for stimulating and helpful discussions during the course of this research.

APPENDIX A: AN R_0 CONSTRAINT FOR INDIVIDUAL-LEVEL MODELS COMPATIBLE WITH THE SI ODE MODEL

We now discuss an R_0 relation which derives from constraining ILMs to be compatible with the SI ODE model at disease invasion. Under the assumption that every infection is uniquely assigned as a secondary infection for either a removed or an infected individual, the following relation holds:

$$N_i(t) = I(t)Q_0(t) + N_r(t)R_0(t), \tag{A1}$$

where $N_r(t) = \int_0^t \mu I(u)du$ is the number of removed individuals at time t, and $N_i(t) = \int_0^t \beta I(u)du$ is the cumulative number of infections that occur in the time interval $(0, t]$. From Eq. (2) we find the prevalence of the infective individuals $I(t)$ and rewrite Eq. (A1) as

$$\frac{(1-\mu/\beta)I(t)}{I(t)-I(0)}Q_0(t) + \frac{\mu}{\beta}R_0(t) = 1, \tag{A2}$$

which is an equation valid for all ILMs that are compatible with the ODE SI model of disease invasion. For large t, we introduce the notation $R_0 \equiv \lim_{t\to\infty} R_0(t)$, and $Q_0 \equiv \lim_{t\to\infty} Q_0(t)$ when the limits exist. We discuss two cases: the nonepidemic and the epidemic cases.

Nonepidemic case. We refer to the situation where the process of infection from individual to individual cannot sustain itself in a susceptible population. This is captured by our paradigm model when $\beta < \mu$. In this case, $I(t) \to 0$ as $t \to \infty$; i.e., no epidemic occurs. In the limit $t \to \infty$, $R_0(t)$ converges to a definite value which is β/μ for *all* ILMs.

Epidemic case. We now assume $\beta > \mu$, and we have a growing epidemic with $I(t) \to \infty$ as $t \to \infty$. As t becomes large, Eq. (A2) takes the following asymptotic form:

$$\left(1 - \frac{\mu}{\beta}\right)Q_0(t) + \frac{\mu}{\beta}R_0(t) = 1, \tag{A3}$$

where convergence of $Q_0(t)$ and $R_0(t)$ as $t \to \infty$ is not guaranteed. From Eq. (A3), notice that $R_0(t)$ cannot exceed β/μ.

APPENDIX B: THE TIME SINCE INFECTION DISTRIBUTION IN THE SI ODE MODEL OF DISEASE INVASION

Here we further assume that removals of individuals in the KMC growing network are independent of each other, and that the removal rate per infectious individual is the constant μ. These assumptions are satisfied by the model described in Sec. III and by the CMJ process. We denote the number of infected individuals that have spent an interval of time between $\tilde{\tau}$ and $\tilde{\tau}+d\tilde{\tau}$ in the I compartment at time t by $\tilde{I}_{\tilde{\tau}}(t)$. Then, we write an equation for $\tilde{I}_{\tilde{\tau}}(t)$ based on the fact that the individuals with time since infection $\tilde{\tau}+dt$ at time $t+dt$ are the individuals with time since infection $\tilde{\tau}$ at time t minus those who have been removed in the dt time interval

$$\tilde{I}_{\tilde{\tau}+dt}(t+dt) = \tilde{I}_{\tilde{\tau}}(t)(1-\mu dt), \tag{B1}$$

which leads to

$$\frac{\partial \tilde{I}_{\tilde{\tau}}(t)}{\partial t} + \frac{\partial \tilde{I}_{\tilde{\tau}}(t)}{\partial \tilde{\tau}} + \mu \tilde{I}_{\tilde{\tau}}(t) = 0. \tag{B2}$$

The solution of Eq. (B2) must satisfy $\int_0^\infty \tilde{I}_{\tilde{\tau}} d\tilde{\tau} = I(t)$, and, since all newly infected individuals that join the infected compartment have $\tilde{\tau}=0$, we also must have $\tilde{I}_0(t) = \beta I(t)$. Such a solution of (B2) is $\tilde{I}_{\tilde{\tau}}(t) = \beta I(0) e^{(\beta-\mu)t} e^{-\beta\tilde{\tau}}$. This solution is approached asymptotically as $t \to \infty$ irrespective of the initial conditions, and leads to the time since infection distribution given by

$$\rho = \tilde{I}_{\tilde{\tau}}(t)/I(t) = \beta e^{-\beta\tilde{\tau}}. \tag{B3}$$

Note that the average time since infection of the population in the infected compartment is β^{-1}.

APPENDIX C: ALTERNATE DERIVATION OF $R_0 = 1$

An alternate explanation of the fact that $R_0 = 1$ for this ILM is as follows. The flow of newly infected individuals is $\beta I(t)$. Thus, the flow per already infected individual is β. Since the removed individuals are randomly sampled from the infectious individuals, the average length of the infectious period equals the time expectation of the infectious period. The average length of the infectious period over the infected individuals is $\langle\tilde{\tau}\rangle = \beta^{-1}$ (see Appendix B), and thus we have $R_0 = \beta\langle\tau\rangle = 1$. In contrast, in the case of the CMJ branching process, the time expectation of the infectious period is μ^{-1}, and thus $R_0^{\mathrm{CMJ}} = \beta/\mu$.

[1] G. MacDonald, Trop. Dis. Bull. **49**, 813 (1952).

[2] K. S. Crump and C. J. Mode, J. Math. Anal. Appl. **24**, 494 (1968).

[3] K. S. Crump and C. J. Mode, J. Math. Anal. Appl. **25**, 8 (1969).

[4] C. J. Mode, *Multitype Branching Processes: Theory and Applications* (American Elsevier, New York, 1971).

[5] C. J. Mode, *Stochastic Processes in Demography and Their Computer Implementation* (Springer-Verlag, Berlin, 1985).

[6] C. J. Mode and C. K. Sleeman, *Stochastic Processes in Epidemiology* (World Scientific, Singapore, 2003).

[7] P. Jagers, Skandinavisk Aktuarietidskift **52**, 84 (1969).

[8] P. Jagers, *Branching Processes with Biological Applications* (John Wiley and Sons, London, 1975).
[9] N. Masuda, N. Konno, and K. Aihara, Phys. Rev. E **69**, 031917 (2004).
[10] C. Moore and M. E. J. Newman, Phys. Rev. E **61**, 5678 (2000).
[11] M. E. J. Newman, I. Jensen, and R. M. Ziff, Phys. Rev. E **65**, 021904 (2002).
[12] M. E. J. Newman, Phys. Rev. E **66**, 016128 (2002).
[13] M. E. J. Newman, Phys. Rev. E **68**, 026121 (2003).
[14] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).
[15] R. Pastor-Satorras and A. Vespignani, Phys. Rev. E **63**, 066117 (2001).
[16] C. P. Warren, L. M. Sander, and I. M. Sokolov, Phys. Rev. E **66**, 056105 (2002).
[17] M. Boguñá, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. Lett. **90**, 028701 (2003).
[18] V. M. Eguíluz and K. Klemm, Phys. Rev. Lett. **89**, 108701 (2002).
[19] M. Kuperman and G. Abramson, Phys. Rev. Lett. **86**, 2909 (2001).
[20] J. M. Read and M. J. Keeling, Proc. R. Soc. London, Ser. B **270**, 699 (2003).
[21] M. Keeling and G. Grenfell, J. Theor. Biol. **203**, 51 (2000).
[22] M. Keeling, Proc. R. Soc. London, Ser. B **266**, 859 (1999).
[23] R. M. May and A. L. Lloyd, Phys. Rev. E **64**, 066112 (2001).
[24] R. M. Anderson and R. M. May, *Infectious Diseases of Humans* (Oxford University Press, Oxford, 1992), pp. 17–19.
[25] S. M. Blower and T. Chou, Nat. Med. **10**, 1111 (2004).
[26] S. M. Blower, A. N. Aschenbach, H. B. Gershengorn, and J. O. Kahn, Nat. Med. **7**, 1016 (2001).
[27] K. Dietz, Stat. Methods Med. Res. **2**, 23 (1993).
[28] J. A. P. Heesterbeek, Acta Biotheor. **50**, 189 (2002).
[29] O. Diekmann, J. A. P. Heesterbeek, and J. A. J. Metz, J. Math. Biol. **35**, 503 (1990).
[30] O. Diekmann and J. A. P. Heesterbeek, *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation* (Wiley, New York, 2000).
[31] P. Van den Driessche and J. Watmough, Math. Biosci. **180**, 29 (2002).
[32] P. P. Potterat, L. Phillips-Plummer, S. Q. Muth, R. B. Rothenberg, D. E. Woodhouse, T. S. Maldonado-Long, H. P. Zimmerman, and J. B. Muth, Sex Transm. Infect. **78** Suppl. I, i159 (2002).
[33] P. P. Potterat, S. Q. Muth, R. B. Rothenberg, H. Zimmerman-Rogers, D. L. Green, J. E. Taylor, M. S. Bonney, and H. A. White, Sex Transm. Infect. **78** Suppl. I, i152 (2002).
[34] D. T. Gillespie, J. Comput. Phys. **22**, 403 (1976).
[35] *Morphological Organizations in Epitaxial Growth and Removal*, edited by Z. Zheng and M. G. Lagally, Directions in Condensed Matter Physics Vol. 14 (World Scientific, Singapore, 1999).
[36] A.-L. Barabási, R. Albert, and H. Jeong, Physica A **272**, 173 (1999).
[37] A.-L. Barabási, R. Albert, and H. Jeong, Physica A **281**, 69 (2000).
[38] A.-L. Barabási, E. Ravasz, and T. Vicsek, Physica A **299**, 559 (2001).
[39] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, Phys. Rev. Lett. **85**, 4633 (2000).
[40] S. N. Dorogovtsev, Phys. Rev. E **67**, 045102(R) (2003).
[41] K. Klemm and V. M. Eguíluz, Phys. Rev. E **65**, 057102 (2002).
[42] P. L. Krapivsky, G. J. Rodgers, and S. Redner, Phys. Rev. Lett. **86**, 5401 (2001).
[43] P. L. Krapivsky and S. Redner, J. Phys. A **35**, 9517 (2002).
[44] J. H. Jones and M. S. Handcock, Proc. R. Soc. London, Ser. B **270**, 1123 (2003).
[45] R. Breban, R. Vardavas and S. Blower (unpublished).
[46] A. L. Lloyd, Theor Popul. Biol. **65**, 49 (2004).
[47] N. T. J. Bailey, *The Elements of Stochastic Processes with Applications to the Natural Sciences* (John Wiley & Sons, New York, 1990).
[48] H. M. Taylor and S. Karlin, *An Introduction to Stochastic Modeling* (Academic Press, London, 1994), Chap. 5.