J. Math. Biol. (2011) 62:143–164
DOI 10.1007/s00285-010-0331-2

Mathematical Biology

Effective degree network disease models

Jennifer Lindquist · Junling Ma · P. van den Driessche · Frederick H. Willeboordse

Received: 12 August 2009 / Revised: 28 January 2010 / Published online: 24 February 2010

Abstract An effective degree approach to modeling the spread of infectious diseases on a network is introduced and applied to a disease that confers no immunity (a Susceptible-Infectious-Susceptible model, abbreviated as SIS) and to a disease that confers permanent immunity (a Susceptible-Infectious-Recovered model, abbreviated as SIR). Each model is formulated as a large system of ordinary differential equations that keeps track of the number of susceptible and infectious neighbors of an individual. From numerical simulations, these effective degree models are found to be in excellent agreement with the corresponding stochastic processes of the network on a random graph, in that they capture the initial exponential growth rates, the endemic equilibrium of an invading disease for the SIS model, and the epidemic peak for the SIR model. For each of these effective degree models, a formula for the disease threshold condition is derived. The threshold parameter for the SIS model is shown to be larger than that derived from percolation theory for a model with the same disease and network parameters, and consequently a disease may be able to invade with lower transmission than predicted by percolation theory. For the SIR model, the threshold condition is equal to that predicted by percolation theory. Thus unlike the classical

J. Lindquist · J. Ma (✉) · P. van den Driessche
Department of Mathematics and Statistics, University of Victoria, Victoria, BC, Canada
e-mail: jma@math.uvic.ca

J. Lindquist
e-mail: jenl@uvic.ca

P. van den Driessche
e-mail: pvdd@math.uvic.ca

F. H. Willeboordse
Department of Physics, National University of Singapore, Singapore, Singapore
e-mail: willeboordse@yahoo.com

homogeneous mixing disease models, the SIS and SIR effective degree models have different disease threshold conditions.

Keywords Network · SIS disease model · SIR disease model · Basic reproduction number

Mathematics Subject Classification (2000) 92D30

1 Introduction

Traditional compartmental epidemic models usually assume that the population is homogeneously mixed, that is, each individual has the same chance of meeting any other individual in the population. This assumption, however, is not realistic. In addition to any spatial heterogeneity, the social structure of the population affects the probability of contact for each pair of individuals. For example, school-age children meet more frequently in class than out of class; this accounts for a major portion of the seasonality observed in measles incidence (Fine and Clarkson 1982). A contact network is a more realistic model of the population. In such a model, each node of the network corresponds to an individual, and each link represents possible contacts between its two nodes; these two nodes are neighbors of each other. The contact rate of a node is then proportional to its degree (the number of neighbors).

In this paper, we concentrate on the susceptible-infectious-susceptible (SIS) and susceptible-infectious-recovered (SIR) models, which assume that susceptibles immediately become infectious once infected, and the infectious individuals become either fully susceptible (SIS) or fully immune (SIR) once recovered. These models are the basic building blocks for more complex disease transmission models, that include, for example, other disease states, multiple groups and age structure.

1.1 Network disease models

Several network disease models have been developed in the literature. Newman (2002) used percolation theory to study the threshold condition and final size of an epidemic for diseases on a contact network. Consider an SIR model on a contact network, in which susceptibles are infected by neighboring infectious individuals with a per link transmission rate β, and infectious individuals recover to full immunity at a rate γ. This model implicitly assumes that the transmission process is Poisson, and that the infectious period is exponentially distributed. The condition necessary for invasion of the population by the disease (disease threshold condition) is that the basic reproduction number $\mathcal{R}_0$, i.e., the expected number of secondary infections caused by a typical infectious individual introduced in a completely susceptible population, is greater than one. Newman's formulation of the SIR disease threshold condition can be written as

$$\mathcal{R}_0 = \frac{\beta}{\beta+\gamma}\frac{\langle k(k-1)\rangle}{\langle k\rangle} = \frac{\beta}{\beta+\gamma}\left(\langle k\rangle - 1 + \frac{\mathrm{Var}\,[k]}{\langle k\rangle}\right) > 1, \tag{1}$$

where $\langle k \rangle$ and Var $[k]$ are the mean and variance of the contact network degree distribution; see, for example, Lloyd and Valeika (2007). Note that $k-1$ is the maximum number of susceptible neighbors that a non-initial infectious individual with degree k may have. The factor $\frac{\beta}{\beta+\gamma}$ is the probability that a link connecting an infectious node and a susceptible node propagates an infection.

One limitation of percolation theory is that it lacks a description of disease dynamics. For example, there is no temporal tracking of disease incidence. To model such variables, a dynamical model portraying disease spread on a contact network may be used. Pastor-Satorras and Vespignani (2001a,b) pioneered this field with the following SIS model (i.e., infectious nodes revert to a susceptible state once recovered, with no acquired immunity):

$$I_k' = \beta k \Theta (N_k - I_k) - \gamma I_k, \quad k = 1, 2, \ldots, M\,, \tag{2}$$

where N_k is the number of nodes with degree k, M is the maximum degree of a node, and $S_k = N_k - I_k$ and I_k are the number of susceptible and infectious individuals with degree k, respectively. The first term on the right hand side of (2) describes the rate of new infections for S_k (thus entering I_k), and the second term represents the recovery of nodes in I_k. The quantity

$$\Theta = \frac{\sum_{i=1}^{M} i I_i}{\sum_{i=1}^{M} i N_i} \tag{3}$$

is the probability that a randomly selected neighbor of a given node is infectious. It follows that $k\Theta$ is the expected number of infectious neighbors of a susceptible node of degree k. Pastor-Satorras and Vespignani (2002) showed that according to this model the disease can invade the population if and only if

$$\mathcal{R}_0 = \frac{\beta}{\gamma} \frac{\langle k^2 \rangle}{\langle k \rangle} = \frac{\beta}{\gamma} \left(\langle k \rangle + \frac{\text{Var}\,[k]}{\langle k \rangle} \right) > 1. \tag{4}$$

The classic homogeneous mixing SIS model (see, for example, Brauer 2008, p. 52), is a special case of model (2) where the contact network is a large complete graph of N nodes. This gives $\langle k \rangle = N - 1 \approx N$, and Var $[k] = 0$; the disease threshold condition given in (4) is then

$$\mathcal{R}_0 = \frac{\beta N}{\gamma} > 1. \tag{5}$$

This is the same disease threshold condition found in traditional homogeneous mixing SIR models (see, for example, Brauer 2008, p. 27).

There are several modifications to the Pastor–Satorras and Vespignani model. For example, Kiss et al. (2006) specified SIR type transmission both through global homogeneous mixing and via a contact network. Considering the special case, in which the

homogeneous mixing is ignored, the Kiss et al. model becomes

$$S_k' = -\beta k \Theta S_k \,, \tag{6a}$$
$$I_k' = \beta k \Theta S_k - \gamma I_k \,, \tag{6b}$$

where Θ in (3) is replaced by

$$\Theta = \frac{\sum_{i=1}^{M} (i-1) I_i}{\sum_{i=1}^{M} i N_i}. \tag{6c}$$

Note that because a non-initial infectious individual must have been infected by a neighbor, a member of the I_k class may transmit to at most $k-1$ individuals. This modified Kiss et al. model (6) results in the disease threshold condition (see Ball and Neal 2008, Sect. 5)

$$\mathcal{R}_0 = \frac{\beta}{\gamma} \left(\langle k \rangle - 1 + \frac{\mathrm{Var}\,[k]}{\langle k \rangle} \right) > 1. \tag{7}$$

Ball and Neal (2008, Sect. 3) used a branching process approximation, adapted from Ball (1983) and Neal (2007), to the early stage of an SIR epidemic with two levels of mixing. They derived the threshold condition which, when ignoring global transmission, is identical to the percolation theory threshold (1). In addition, Ball and Neal (2008, Sect. 5) formulated a deterministic approximation to a stochastic SIR epidemic with global and local transmission. They started with individuals having unpaired half-links, and constructed a contact network while the epidemic evolved. The individuals in this model were classified by the number of unpaired half-links of a node, which Ball and Neal called the effective degree of a node.

Volz (2008) introduced a dynamical probability generating function (PGF) disease model, which has the advantage of using only three differential equations. This approach dynamically describes the SIR disease process, while accounting for the network structure. The model incorporates the probabilities that a neighbor of a susceptible individual is susceptible or infectious. The SIR disease threshold condition of this model is shown to be equivalent to the condition given by percolation theory (1).

To compare the contact network formulations with homogeneous mixing models, note that βN is the per capita transmission rate of an infectious individual in a completely susceptible homogeneously mixed population, and that $\beta \langle k \rangle$ is the average per capita transmission rate of such an individual in the network model (2). For a given disease context the average per capita transmission rate of an infectious individual is assumed to be a fixed number. Hence $\mathrm{Var}\,[k] \geq 0$ implies that $\mathcal{R}_0$ given by (4) is larger than that in (5). Consequently, for a given disease, the Pastor-Satorras and Vespignani model predicts faster spread of the disease than the homogeneous model.

There are also differences between actual networks and the models used to approximate them. For example, the Pastor-Satorras and Vespignani model and its variations are intrinsically heterogeneous mixing models. In such models, individuals randomly

interact with other members of the population with probabilities based on the degree of the node contacted. Hence, an infectious individual with one transmittable link can transmit to multiple susceptibles before it recovers. This cannot happen on a contact network: in the SIS case, once the susceptible neighbor is infected, it cannot be infected again until it recovers and becomes newly susceptible. In the SIR case, this means that the infectious individual with one transmittable link may transmit at most once; furthermore, a degree one infectious node can only transmit if it is an index case.

Bansal et al. (2007) compared stochastic SIR disease spread on various random networks with the predictions of four network models (heterogeneous mixing, PGF, percolation theory, and pair approximation) and the classic SIR homogeneous mixing model. The inclusion of percolation theory dictates that only comparisons of epidemic final size or related quantities can be considered. The authors showed that not only do the various models differ from one another and the underlying stochastic process, but also that the degree distribution of the network has a significant effect on a model's predictions and accuracy with respect to the final size of an SIR epidemic.

To examine the dynamical accuracy of existing models, we performed 100 stochastic simulations (see Appendix A for methods) of an SIS disease process on the same randomly generated contact network, and compared these with the predictions of the Pastor–Satorras and Vespignani model (2), and the classic homogeneous mixing model. Figure 1 shows that on a random contact network disease in a stochastic model spreads slower than in either the homogeneous mixing model or the Pastor–Satorras and Vespignani model. The simulations also indicate that the disease threshold condition given by percolation theory is inaccurate in the case of an SIS epidemic process. Using the parameter values of the simulations ($\beta = 0.05$, $\gamma = 0.2$), the percolation theory threshold parameter given in (1) yields $\mathcal{R}_0 = 0.95 < 1$, indicating that the disease should die out. However, our simulations clearly show that the disease becomes established and reaches an endemic equilibrium.

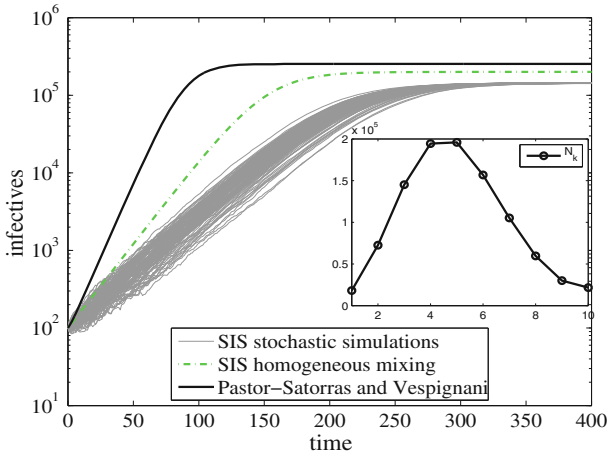

Fig. 1 100 runs of stochastic simulations of the network SIS model on a random graph with $\beta = 0.05$, $\gamma = 0.2$, $I_0 = 100$ and the degree distribution (shown in the *inset*) $N_1 = 18,118$, $N_2 = 72,536$, $N_3 = 145,222$, $N_4 = 194,589$, $N_5 = 195,962$, $N_6 = 156,857$, $N_7 = 105,280$, $N_8 = 59,713$, $N_9 = 30,066$, $N_{10} = 21,657$. This corresponds to $N = 10^6$, $\langle k \rangle = 5$, and $M = 10$

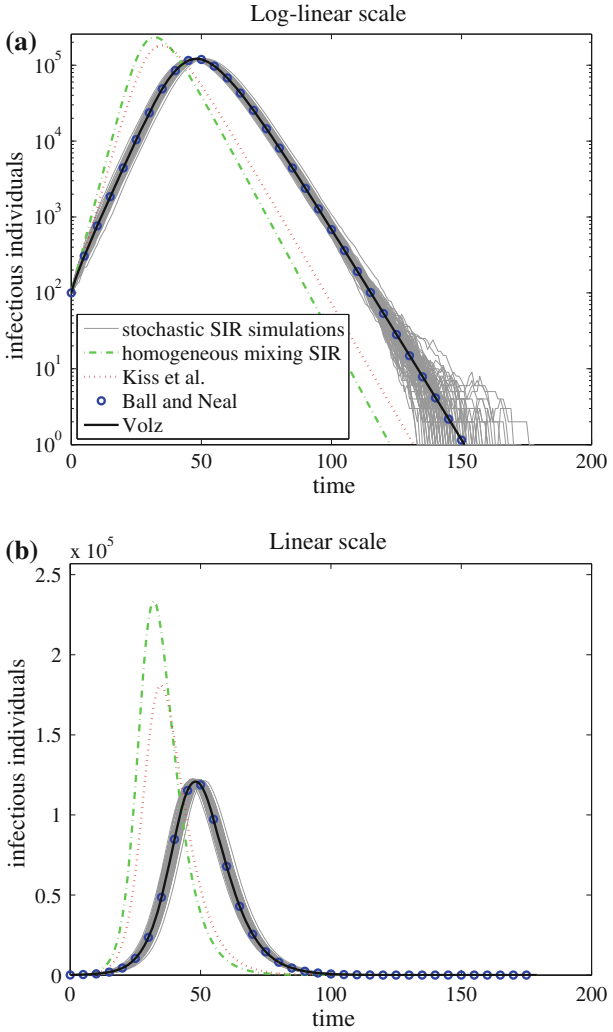

Fig. 2 100 runs of stochastic simulations of the network SIR model on a random graph, shown in both log-linear (highlighting the *exponential growth phase*) and linear (highlighting the *peak*) scales. In this figure, $\beta = 0.1$, $\gamma = 0.2$, $I_0 = 100$ and the contact network is identical to that used in Fig. 1

Figure 2 shows the results of 100 stochastic simulations of an SIR epidemic on the same contact network as used for Fig. 1, compared with predictions of the homogeneous mixing, the Kiss et al., the Volz, and the Ball and Neal effective degree models. Figure 2a shows that the Kiss et al. and the homogeneous mixing models have an exponential growth rate that is too large. On the other hand, the Volz model and Ball and Neal effective degree model closely match the stochastic simulations over the whole epidemic as can be clearly seen in Fig. 2b.

Thus unlike the classic homogeneous mixing models, network models appear to have different threshold conditions in the SIR and SIS processes. Among the currently

available models, the Volz model and the Ball and Neal model appear to best approximate the stochastic SIR epidemic process on a contact network. However, these two models cannot easily be extended to SIS type diseases. In order to understand this discrepancy between the SIR and SIS epidemic thresholds, we develop a new class of models. We call these effective degree models following (Ball and Neal 2008). However, our model setup is different from that of Ball and Neal.

1.2 Effective degree models

For the effective degree approach, each node of a disease network is categorized by its disease state, e.g., susceptible (S), infectious (I), or recovered (R), as well as by the number of its neighbors in each disease state. As a consequence, the network is divided into classes representing the state of an individual and that of its neighbors. For example, S_{si} denotes the number of susceptible individuals with s susceptible neighbors and i infectious neighbors. We call s the susceptible degree and i the infectious degree. If recovered individuals acquire life long immunity, then these individuals do not participate further in disease transmission, and thus can be removed from the network together with their links. In this case the model does not need to keep track of the R state, and nodes change their degree as their neighbors recover. When a node changes state, it moves to the corresponding class with the same subscripts, because its neighbors do not change their infection state. However, its neighbors do change their subscripts while keeping their disease state. To illustrate, suppose a node is a member of S_{si}. If this node is infected, then it moves to class I_{si}. In the SIS case, the recovery of a neighbor would result in the S_{si} node moving to the $S_{s+1,i-1}$ class, whereas under the SIR model this would result in this node entering the $S_{s,i-1}$ class. The infection of a neighbor moves the S_{si} node to the $S_{s-1,i+1}$ class in both the SIS and the SIR models.

In Sect. 2, we apply this effective degree approach to an SIS model, and derive the disease threshold condition. In Sect. 3, we apply the same approach to an SIR model. Our results show that the disease threshold parameter for the network SIS model is larger than that of the network SIR model. In Sect. 4, we discuss the differences found between these disease threshold conditions.

2 SIS model

In an SIS model there are two disease states, susceptible (S) and infectious (I). Such a model is appropriate for many bacterial diseases, for example gonorrhea. Following the notation of Sect. 1 we now specify the SIS effective degree model. We consider a randomly connected network without multiple links, self loops, clustering, or degree correlation.

2.1 SIS model formulation

As introduced in Sect. 1, γ and β denote the per node recovery rate and per link transmission rate, respectively. Following the notation of Sect. 1.2, we denote the number

of susceptible and infectious individuals with s susceptible neighbors and i infectious neighbors by S_{si} and I_{si}, respectively. A susceptible individual in class S_{si} is infected with rate $i\beta$, while an infectious node in I_{si} recovers and enters class S_{si} with rate γ. Because each node in S_{si} has i infectious neighbors, and each neighbor recovers at rate γ, the rate with which nodes in S_{si} leave this class and enter $S_{s+1,i-1}$ due to recovery of a neighbor is $\gamma i S_{si}$.

The infection of a neighbor of a node increases its infectious degree by one, and decreases its susceptible degree by one. The rate of new infections is $\sum_{k=1}^{M}\sum_{j+l=k}\beta l S_{jl}$. These new infections cause their susceptible neighbors to change their effective degree at the rate $\sum_{k=1}^{M}\sum_{j+l=k} j\beta l S_{jl}$. The probability that a susceptible node of S_{si} is a neighbor of a newly infected node is proportional to its susceptible degree (as the newly infected nodes were originally susceptible), that is,

$$\frac{s S_{si}}{\sum_{k=1}^{M}\sum_{j+l=k} j S_{jl}}.$$

Hence, the rate that nodes in S_{si} leave the class and enter $S_{s-1,i+1}$ due to the infection of a neighbor is

$$\frac{\sum_{k=1}^{M}\sum_{j+l=k} j\beta l S_{jl}}{\sum_{k=1}^{M}\sum_{j+l=k} j S_{jl}} s S_{si}.$$

Similarly, the rate that nodes in $S_{s+1,i-1}$ leave the class and enter S_{si} due to the infection of a neighbor is

$$\frac{\sum_{k=1}^{M}\sum_{j+l=k} j\beta l S_{jl}}{\sum_{k=1}^{M}\sum_{j+l=k} j S_{jl}} (s+1) S_{s+1,i-1}\,,$$

and these are the only nodes that enter S_{si} due to the infection of a neighbor.

The rates of leaving and entering I_{si} due to the infection status change of its neighbors are derived similarly. The flow chart of the SIS system is shown in Fig. 3. The system of $\sum_{k=1}^{M} 2(k+1) = M(M+3)$ equations that govern the dynamics of an SIS model on a static contact network is thus

$$\begin{aligned} S'_{si} &= -\beta i S_{si} + \gamma I_{si} + \gamma\left[(i+1)\, S_{s-1,i+1} - i S_{si}\right] \\ &\quad + \frac{\sum_{k=1}^{M}\sum_{j+l=k} j\beta l S_{jl}}{\sum_{k=1}^{M}\sum_{j+l=k} j S_{jl}}\left[(s+1)\, S_{s+1,i-1} - s S_{si}\right], \qquad (8a) \\ I'_{si} &= \beta i S_{si} - \gamma I_{si} + \gamma\left[(i+1)\, I_{s-1,i+1} - i I_{si}\right] \\ &\quad + \frac{\sum_{k=1}^{M}\sum_{j+l=k} \beta l^2 S_{jl}}{\sum_{k=1}^{M}\sum_{j+l=k} j I_{jl}}\left[(s+1)\, I_{s+1,i-1} - s I_{si}\right], \qquad (8b) \end{aligned}$$

for $\{(s,i) : s \geq 0, i \geq 0, s+i \leq M\}$. We call this system the *SIS effective degree model*.

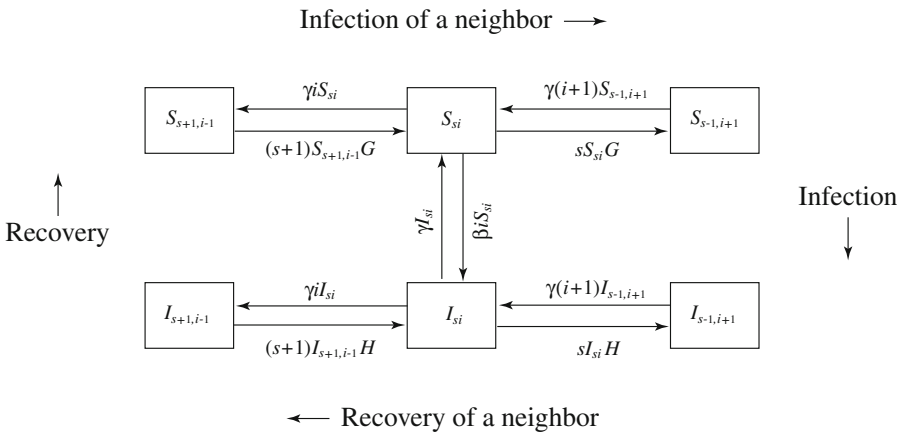

Fig. 3 Flow chart of the SIS effective degree model for the nodes S_{si} and I_{si}, where $G = \frac{\sum_{k=1}^{M}\sum_{j+l=k} j\beta l S_{jl}}{\sum_{k=1}^{M}\sum_{j+l=k} jS_{jl}}$, and $H = \frac{\sum_{k=1}^{M}\sum_{j+l=k} \beta l^2 S_{jl}}{\sum_{k=1}^{M}\sum_{j+l=k} jI_{jl}}$

This model has a disease free equilibrium, namely, $I_{si} = 0$ for all s and i, $S_{si} = 0$ for all $i \geq 1$, and $S_{s0} = N_s$ (the network degree distribution) for all s. Note that

$$\frac{sI_{si}}{\sum_{k=1}^{M}\sum_{j+l=k} jI_{jl}} \leq 1,$$

and that

$$\sum_{k=1}^{M}\sum_{j+l=k} \beta l^2 S_{jl} = 0$$

at the disease free equilibrium, so the last term in (8b) is defined to be zero at the disease free equilibrium.

The SIS effective degree model given by (8) satisfies a balance condition, that is, the total number of infectious neighbors of all susceptible nodes equals the total number of susceptible neighbors of all infectious nodes, because they count the links that connect an infectious node and a susceptible node. Mathematically, this is

$$\sum_{k=1}^{M}\sum_{j+l=k} lS_{jl} = \sum_{k=1}^{M}\sum_{j+l=k} jI_{jl}. \tag{9}$$

Initially, $S_{s0} \leq N_s$ are non-negative and arbitrary, there are some infectious individuals, and the balance condition (9) is satisfied. This balance condition is proved in Appendix B to hold for all $t \geq 0$. Solutions S_{si} and I_{si} to (8) with these initial conditions remain non-negative for all $t \geq 0$.

2.2 Comparison with stochastic simulations

With the stochastic SIS network model specified in Appendix A, we now numerically verify that the effective degree model is a good approximation to the stochastic process. We simulate the SIS epidemic process on the same contact network that is used in the simulations shown in Figs. 1 and 2. The number of initial infections $I_0 = 100$. The recovery rate $\gamma = 0.2$ (mean infectious period is 5), the per link transmission rate $\beta = 0.05$. We simulate the stochastic model for 100 independent realizations. The results show that the SIS effective degree model closely captures the exponential growth rate (see Fig. 4a) and the endemic equilibrium value of the stochastic process (see Fig. 4b).

2.3 SIS disease threshold condition

Applying the balance condition (9) to the denominator in (8b) gives

$$I'_{si} = \beta i S_{si} - \gamma I_{si} + \gamma \left[(i+1)\, I_{s-1,i+1} - i I_{si}\right] + \frac{\sum_{k=1}^{M} \sum_{j+l=k} \beta l^2 S_{jl}}{\sum_{k=1}^{M} \sum_{j+l=k} l S_{jl}} \left[(s+1)\, I_{s+1,i-1} - s I_{si}\right].$$

Near the disease free equilibrium, i.e., when the number of infectious nodes is small, it is unlikely that a susceptible node has more than one infectious neighbor, thus, $S_{si} \approx 0$ for all $i > 1$, giving

$$\frac{\sum_{k=1}^{M} \sum_{j+l=k} \beta l^2 S_{jl}}{\sum_{k=1}^{M} \sum_{j+l=k} l S_{jl}} = \frac{\sum_{j=0}^{M-1} \beta S_{j1}}{\sum_{j=0}^{M-1} S_{j1}} = \beta.$$

Hence, near the disease free equilibrium,

$$I'_{si} = \beta i S_{si} - \gamma I_{si} + \gamma \left[(i+1)\, I_{s-1,i+1} - i I_{si}\right] + \beta \left[(s+1)\, I_{s+1,i-1} - s I_{si}\right]. \tag{10}$$

However, this approximation does not hold when the system evolves away from the disease free equilibrium.

Using (8a) linearized at the disease free equilibrium and (10), the stability of the disease free equilibrium yields a disease threshold condition for the effective degree model. Infection causes individuals in class S_{s0} to move into class $S_{s-1,1}$ at a rate

$$\frac{\sum_{k=1}^{M} \sum_{j+l=k} j\beta l S_{jl}}{\sum_{k=1}^{M} \sum_{j+l=k} j S_{jl}} s S_{s0},$$

and then they can be infected by their infectious neighbor. Other terms in the equations are transfers between classes S_{si} for $i \geq 1$ and I_{si} for all s and i. To use the

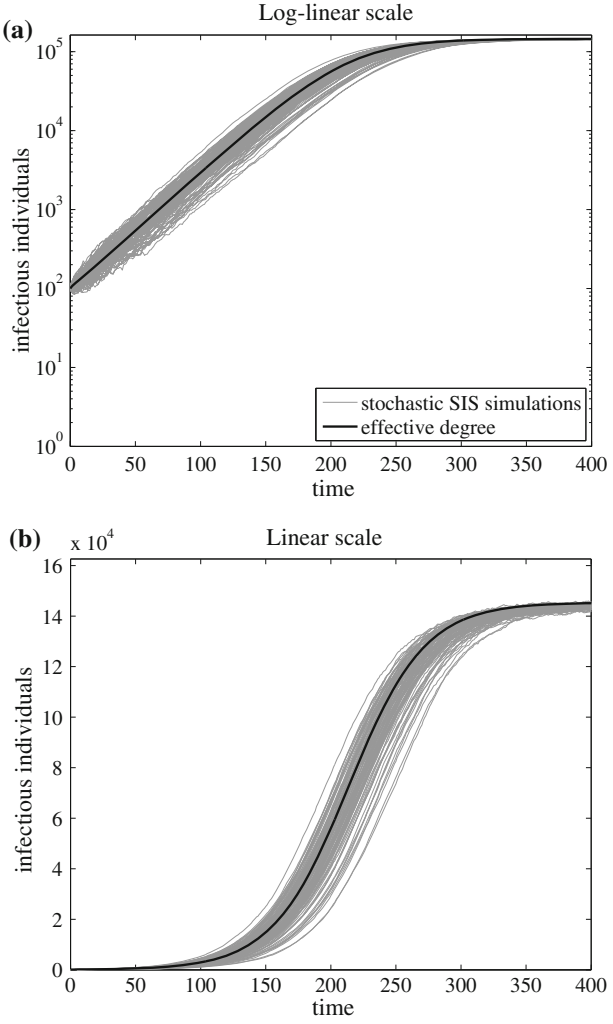

Fig. 4 Comparison of the stochastic SIS and the SIS effective degree models on a random graph, shown in both *log-linear* and *linear scales*. Both models yield the same exponential growth rate of the disease and the same endemic equilibrium value. $\beta = 0.05$, $\gamma = 0.2$, and $I_0 = 100$. The simulation uses the same contact network as in Fig. 1

next generation matrix method (Diekmann and Heesterbeek 2000; van den Driessche and Watmough 2002), we write the Jacobian matrix at the disease free equilibrium as $F - V$, where F is the matrix involving the flow from S_{s0} to $S_{s-1,1}$, and $-V$ is the matrix with transfer terms. For a fixed degree $1 \leq k \leq M$ there are $2(k+1)$ equations from (8a) and (10). But since S_{k0} does not enter into any equation except that of S'_{k0}, these equations can be ignored for the stability calculation. Order the variables for a fixed k with S_{si} equations first, then I_{si} equations arranged lexicographically, and then in increasing values of k; for example, S_{01}, I_{01}, I_{10}; S_{11}, S_{02}, I_{20}, I_{11}, I_{02};. . ., I_{0M}.

Matrix F has rank 1, and is denoted by

$$F = \frac{\beta}{\sum_{k=1}^{M} kS_{k0}} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_M \end{bmatrix} \left[v_1^T, v_2^T, \ldots, v_M^T \right], \tag{11}$$

where u_k is a $(2k+1) \times 1$ vector with kS_{k0} in its first entry and zeros elsewhere, and v_k has the same order as u_k, with its first $(k-1)$ entries equal to $(k-1), 2(k-2), \ldots, s(k-s), \ldots, (k-1)$, and its remaining entries equal to zero. Since the degree of each vertex remains fixed, V is a block diagonal matrix, namely $V = V_1 \bigoplus V_2 \bigoplus \ldots \bigoplus V_M$. Each V_k is of order $(2k+1)$ and depends on β and γ. In addition, each V_k^{-1} exists and is positive, since V_k is a nonsingular M-matrix; see, for example, Varga (1962, Theorem 1.8). Using ρ to denote the spectral radius, the disease threshold is then

$$\mathcal{R}_0 = \rho(FV^{-1}) = \frac{\beta}{\sum_{k=1}^{M} kS_{k0}} \sum_{k=1}^{M} v_k^T V_k^{-1} u_k. \tag{12}$$

If $\mathcal{R}_0 < 1$, then the disease free equilibrium is stable so disease cannot invade; whereas if $\mathcal{R}_0 > 1$, then the disease free equilibrium is unstable so the number of infectious nodes first increases. Simulations show that the disease converges to an endemic equilibrium. For given parameters β, γ, M, and N_k, this formula (12) is easy to implement numerically. For example, with the contact network used in Fig. 1 and $\gamma = 0.2$, the threshold transmission rate $\beta = 0.042$ gives $\mathcal{R}_0 = 1$; we confirmed by stochastic simulations that β is in the interval $(0.0415, 0.0425)$.

We prove in Appendix C that for the same disease and network parameters, the SIS epidemic threshold $\mathcal{R}_0$ as given in (12) is always larger than that of the percolation theory (1). In other words, percolation theory overestimates the transmission rate β required for the invasion of an SIS type disease.

3 SIR model

Many viral diseases, for example, childhood diseases such as measles, confer immunity against reinfection. Thus an SIR model, in which infectious individuals acquire immunity and will not be infected again upon recovery, is appropriate for considering the time evolution of such diseases. We consider the same type of random contact network as in Sect. 2.

3.1 SIR model formulation

To formulate our SIR effective degree model, we use the notation of our SIS effective degree model (Sect. 2.1). Recovered individuals play no further role in the disease transmission process, thus we do not keep track of their numbers, instead, we remove

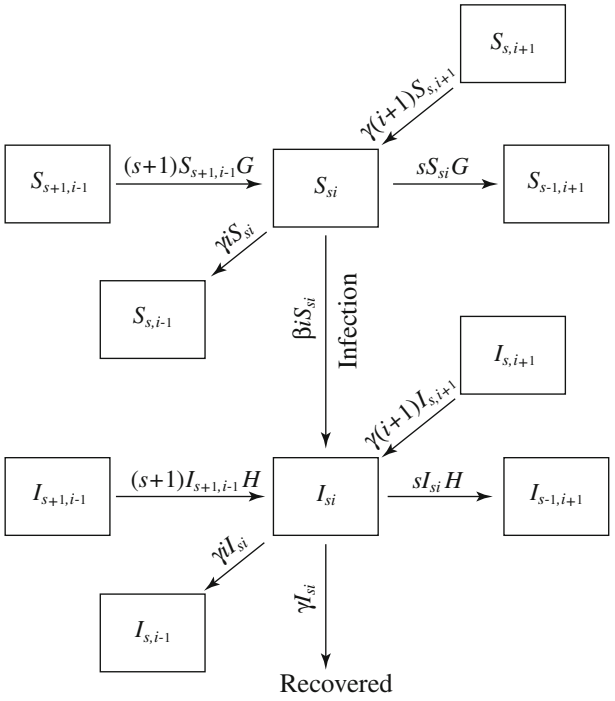

Fig. 5 Flow chart for the nodes S_{si} and I_{si} of the SIR effective degree model, where $G = \frac{\sum_{k=1}^{M}\sum_{j+l=k} j\beta l S_{jl}}{\sum_{k=1}^{M}\sum_{j+l=k} jS_{jl}}$, and $H = \frac{\sum_{k=1}^{M}\sum_{j+l=k} \beta l^2 S_{jl}}{\sum_{k=1}^{M}\sum_{j+l=k} jI_{jl}}$

them from the network. Thus, unlike in our SIS model, the degree $s+i$ of each susceptible or infectious node may decrease as their infectious neighbors recover and are removed from the network.

For a node in class S_{si}, the recovery of one of its neighbors decreases its infectious degree by one while leaving the susceptible degree unchanged, and hence the node moves to class $S_{s,i-1}$. The rate at which such nodes leave S_{si} and enter $S_{s,i-1}$ is given by $\gamma i S_{si}$, as each infectious neighbor recovers with rate γ. Similarly, the rate at which nodes in I_{si} enter $I_{s,i-1}$ is given by $\gamma i I_{si}$. Infectious nodes in I_{si} recover and leave the network with rate γI_{si}. The infection terms for this SIR model are the same as for the SIS model formulated in Sect. 2.1. The flow chart for the SIR system is shown in Fig. 5. Note that since nodes may be removed from the network, it is possible that a node has degree zero, that is, $s+i$ can be zero.

Our SIR effective degree model on a contact network is thus governed by the system of $M(M+3)$ equations:

$$S'_{si} = -\beta i S_{si} + \gamma\left[(i+1)\,S_{s,i+1} - iS_{si}\right] + \frac{\sum_{k=1}^{M}\sum_{j+l=k} j\beta l S_{jl}}{\sum_{k=1}^{M}\sum_{j+l=k} jS_{jl}}\left[(s+1)\,S_{s+1,i-1} - sS_{si}\right], \tag{13a}$$

$$I'_{si} = \beta i S_{si} - \gamma I_{si} + \gamma \left[(i+1) I_{s,i+1} - i I_{si}\right] + \frac{\sum_{k=1}^{M} \sum_{j+l=k} \beta l^2 S_{jl}}{\sum_{k=1}^{M} \sum_{j+l=k} j I_{jl}} \left[(s+1) I_{s+1,i-1} - s I_{si}\right], \tag{13b}$$

for $\{(s,i) : s \geq 0, i \geq 0, s+i \leq M\}$. Initially, $S_{s0} \leq N_s$ are nonnegative and arbitrary, there are some infectious nodes and no recovered ones, and the balance condition (9) holds. Using arguments similar to those in Appendix B, the balance condition can be shown to hold for the SIR model given by (13) for all $t \geq 0$. Solutions to this system with the given initial conditions remain nonnegative and bounded for all $t \geq 0$.

Summing (13a) and using the balance condition (9) gives

$$S' = \sum_{k=0}^{M} \sum_{s+i=k} S'_{si} = -\beta \sum_{k=0}^{M} \sum_{s+i=k} s I_{si} \leq 0.$$

Thus S is decreasing and bounded below, so $\lim_{t\to\infty} S' = 0$. With $I' = \sum_{k=0}^{M} \sum_{s+i=k} I'_{si}$, summing (13b) gives

$$I' = -S' - I = -I$$

as $t \to \infty$. Hence, $I \to 0$ as $t \to \infty$. That is, the disease will eventually die out. In fact, I may decrease to zero (no epidemic) or first increase to a maximum and then decrease to zero (an epidemic).

3.2 Comparison with stochastic simulations

We now numerically compare the SIR effective degree model with the stochastic SIR network model specified in Appendix A. We numerically evaluate the SIR effective degree model (13) on the same contact network used in the simulations shown in Figs. 1, 2, and 4. The transmission rate $\beta = 0.1$, and the recovery rate $\gamma = 0.2$. Figure 6 shows that the SIR effective degree model closely captures the epidemic curves of the stochastic SIR process.

From Fig. 6, the final size of the epidemic is not immediately apparent; whereas the endemic equilibrium values of the SIS effective degree model and the stochastic SIS simulations can be read off from Fig. 4. Thus, Fig. 7 is given to compare the final sizes of the SIR effective degree model with other SIR models and our stochastic SIR simulations. As can be seen, the final size of the SIR effective degree model provides the best match with the stochastic simulations among the models compared, although the difference with Ball and Neal is so small as to be insignificant while the difference with Voltz is only very minor. Consequently, with respect to final size, the SIR effective degree model, the Ball and Neal model and the Voltz model can be considered to perform equally well and their relative merits are determined by other factors such as simplicity or adaptability to different network topologies.

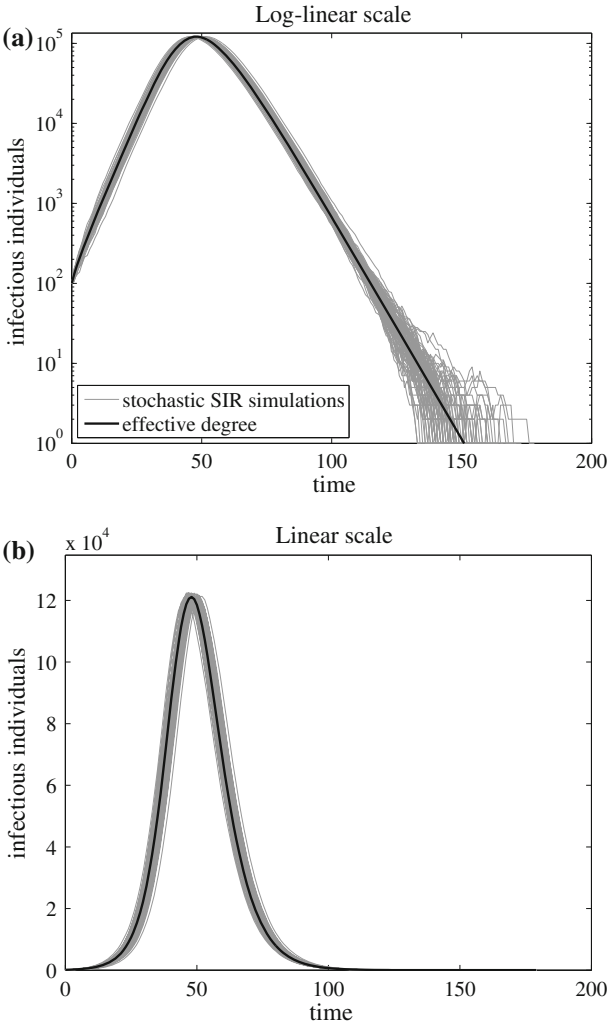

Fig. 6 Comparison of the stochastic SIR and the SIR effective degree models on a random graph, shown in both *log-linear* and *linear scales*. The effective degree SIR model has the same exponential growth rate and peak as those of the stochastic SIR model. $\beta = 0.1$, $\gamma = 0.2$, and $I_0 = 100$. The simulation uses the same contact network as in Fig. 1

3.3 SIR disease threshold condition

The occurrence or not of an epidemic for a given set of parameter values depends on the disease threshold condition. The disease free equilibrium of the SIR model given by (13) has the same form as that of the SIS model in Sect. 2.1. Thus, by using the same approximation as in the SIS model (Sect. 2.3) near the disease free equilibrium,

$$I'_{si} = \beta i S_{si} - \gamma I_{si} + \gamma \left[(i+1)\, I_{s,i+1} - i I_{si}\right] + \beta \left[(s+1)\, I_{s+1,i-1} - s I_{si}\right].$$

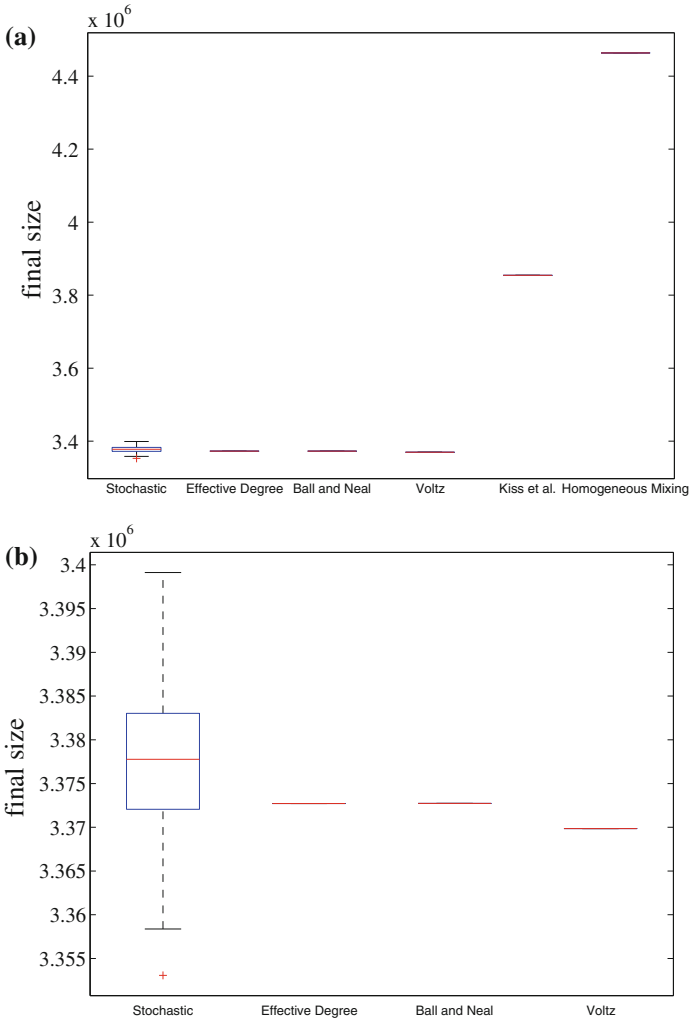

Fig. 7 **a** Comparison of the final sizes of the SIR epidemic produced by the stochastic simulations and the various ordinary differential equation models on the same random graph as in Fig. 1 with identical initial conditions. The parameter values $\beta = 0.1$, $\gamma = 0.2$, and $I_0 = 100$ are used in all models. The final sizes of the 100 stochastic simulations are represented by a *box-whisker plot*, in which the box represents the 50% confidence interval, the whiskers represent the extent of the points that are not outliers, and the + represents the outlier. **b** The region containing the final sizes for the stochastic, effective degree, Ball and Neal, and Voltz models is enlarged for better visibility of the details

Using this and (13a) linearized at the disease free equilibrium, we proceed as in Sect. 2.3 by writing the Jacobian matrix as $F - V$ and using this same ordering of variables. For this model, the F matrix is the same as in (11) except with an extra $u_0 = v_0 = 0$, whereas the V matrix is block upper triangular with diagonal blocks V_{ii}, $i = 0, \ldots, M$, and superdiagonal blocks $V_{i,i+1}$, $i = 0, \ldots, M-1$. Each V_{ii} is lower triangular with positive diagonal entries and each $V_{i,i+1}$ has every entry in its

first column equal to zero. Thus,

$$\rho(FV^{-1}) = \sum_{i=0}^{M} F_{ii} V_{ii}^{-1}$$

whereas

$$F_{ii} = \frac{\beta}{\sum_{k=0}^{M} k S_{k0}} u_i v_i^T ,$$

and the (1, 1) entry of $V_{ii}^{-1} = 1/(\beta+\gamma)$. Thus, given $S_{k0} = N_k$, the disease threshold parameter is given by

$$\mathcal{R}_0 = \rho(FV^{-1}) = \frac{\beta}{\beta+\gamma} \frac{\sum_{k=0}^{M} k(k-1) S_{k0}}{\sum_{k=0}^{M} k S_{k0}} = \frac{\beta}{\beta+\gamma} \frac{\langle k(k-1) \rangle}{\langle k \rangle}. \tag{14}$$

If $\mathcal{R}_0 < 1$, then the disease free equilibrium is stable so disease cannot invade; whereas if $\mathcal{R}_0 > 1$, then the disease free equilibrium is unstable so the number of infectious nodes first increases giving an epidemic, and then decreases to zero.

This disease threshold parameter for the SIR effective degree model (14) is equal to that of percolation theory (1). Interestingly, it is not equal to that of the SIS effective degree model (12) with the same parameters. In fact, the threshold parameter of our SIR model is less than that of our SIS model (see Appendix C).

4 Discussion

We formulate network SIS and SIR effective degree models as systems of ordinary differential equations that keep track of the number of susceptible and infectious neighbors of a node. Numerical simulations show that these models closely capture the time evolution of the disease, notably the exponential growth rate. Additionally, the SIS model closely predicts the endemic equilibrium (see Fig. 4b), while the SIR model captures the peak of the epidemic curve (see Fig. 6b).

The disease threshold parameter of the SIS effective degree model can be computed numerically given the network degree distribution, the disease transmission rate and the recovery rate. The resulting threshold condition requires a smaller transmission rate β for the invasion of the disease than the percolation theory threshold condition. On the other hand, the disease threshold condition of the SIR effective degree model is the same as the percolation theory threshold condition, as also predicted by Volz (2008). However, importantly, the SIS and SIR effective degree models do not have the same threshold condition. Indeed, the threshold parameter of the SIS model is larger than that of the SIR model. This is due to the fact that when disease transmission progresses and infectious nodes recover, the recovered nodes block the transmission paths in the SIR model whereas in the SIS model these paths are freed when infectious nodes recover. This is reflected by the degree change of the nodes in our SIR model as

the epidemic progresses. Our result contrasts with the traditional homogeneous and heterogeneous mixing SIS and SIR models, for which the disease threshold conditions are identical. This is because on a static contact network the neighbors of an infectious node are fixed, and thus the infection and recovery of its neighbors reduce the number of its susceptible neighbors, whereas for the traditional homogeneous and heterogeneous mixing models, the recovery of one node has little effect on the availability of susceptibles.

For the Pastor–Satorras and Vespignani model (2), it is assumed that the probability that a neighbor is infectious does not depend on the degree. To compare our SIS effective degree model with (2), define $S_k = \sum_{s+i=k} S_{si}$, $I_k = \sum_{s+i=k} I_{si}$, and

$$\Theta_k = \frac{\sum_{s+i=k} i S_{si}}{k S_k}.$$

Note that Θ_k is the probability that a neighbor of a degree k node is infectious. Then, from (8b),

$$I_k' = \beta k \Theta_k (N_k - S_k) - \gamma I_k. \tag{15}$$

The difference between (15) and the Pastor-Satorras and Vespignani model is that in our model the probability Θ_k depends on the degree k.

For the effective degree models, we assume that the probability that a node is a neighbor of a given node is a constant, hence the connectivity variation can be captured by the network distribution only. Thus, our models may not be applicable to small world networks, because in such networks most links are not random. Our models describe disease spread on a static randomly connected contact network with no clusting or degree correlations, and appear to be excellent candidates as building blocks to construct models involving dynamic networks.

Acknowledgments This research is partially funded by NSERC Discovery Grants (JM, PvdD), MITACS (PvdD), and CIHR (JM). The authors thank the two anonymous referees for helpful comments and suggestions.

Appendix A: The stochastic simulations

We construct as follows an Erdős–Rényi type random graph (Erdős and Rényi 1959) on $N = 10^6$ nodes with maximum allowable degree $M = 10$. Firstly, each of the nodes is assigned a link to a randomly selected node with degree less than M. Secondly, two different nodes each with degree less than M that are not neighbors are randomly selected, and a link between them is added. This second step is repeated until the average degree is equal to $\langle k \rangle = 5$. This random graph, which has no self-loops, multiple links, clustering, or degree-correlation, is used as the contact network in all the stochastic simulations.

In the stochastic SIS network model, each node in the contact network has one of two states: susceptible (S) or infectious (I). For each infectious node, its state changes

to S with rate γ. For each susceptible node, its state changes to I with rate $i\beta$, where i is the number of infectious neighbors of that node.

In the stochastic SIR network model, each node has one of three states: S and I as in the SIS model, and recovered (R). Each infectious node recovers and changes its state to R with rate γ. The transmission process is defined to be the same as in the SIS model.

We simulate the stochastic network SIS and SIR processes with the Gillespie algorithm (Gillespie 1976, 1977). This algorithm avoids the discretization of time steps and operates in continuous time. Gillespie (1976, 1977) proved that the algorithm is exact, which means that the simulations follow a distribution that satisfies the master equation of the underlying stochastic process. To implement this, at $t = 0$ we randomly pick I_0 nodes, set their state to I, and set all the other nodes to S. For each node, a random event time is generated by adding to the current time an exponentially distributed time step with rate either $i\beta$ (if the node is susceptible) or γ (if the node is infectious). Then, repeatedly, the node with the smallest time step is chosen, the current time is updated to the event time of the chosen node, and the state of the chosen node is changed accordingly. The event time of the chosen node and all of its neighbors are then updated. This repetition continues until the given simulation time is reached or there are no infectious nodes in the contact network.

Appendix B: Proof of the balance condition

To prove the balance condition (9), compute from (8a)

$$\begin{aligned}\sum_{k=1}^{M}\sum_{j+l=k} lS'_{jl} &= -\sum_{k=1}^{M}\sum_{j+l=k} l\beta l S_{jl} + \sum_{k=1}^{M}\sum_{j+l=k} \gamma l I_{jl} \\ &\quad +\gamma\sum_{k=1}^{M}\sum_{j+l=k}\left[l(l+1)S_{j-1,l+1} - l^2 S_{jl}\right] \\ &\quad +\left[\sum_{k=1}^{M}\sum_{j+l=k} j\beta l S_{jl}\right]\frac{\sum_{k=1}^{M}\sum_{j+l=k}\left[l(j+1)S_{j+1,l+1} - ljS_{jl}\right]}{\sum_{k=1}^{M}\sum_{j+l=k} jS_{jl}}.\end{aligned}$$

Since

$$\begin{aligned}\sum_{k=1}^{M}\sum_{j+l=k} l(l+1)S_{j-1,l+1} &= \sum_{k=1}^{M}\sum_{j+l=k}(l+1)^2 S_{j-1,l+1} - \sum_{k=1}^{M}\sum_{j+l=k}(l+1)S_{j-1,l+1} \\ &= \sum_{k=1}^{M}\sum_{j+l=k} l^2 S_{jl} - \sum_{k=1}^{M}\sum_{j+l=k} lS_{jl},\end{aligned}$$

it follows that

$$\sum_{k=1}^{M}\sum_{j+l=k}\Big[l(l+1)S_{j-1,l+1}-l^2S_{jl}\Big]=-\sum_{k=1}^{M}\sum_{j+l=k}lS_{jl}.$$

Similarly,

$$\begin{aligned}\sum_{k=1}^{M}\sum_{j+l=k}l(j+1)S_{j+1,l-1}&=\sum_{k=1}^{M}\sum_{j+l=k}(l-1)(j+1)S_{j+1,l-1}\\&\quad+\sum_{k=1}^{M}\sum_{j+l=k}(j+1)S_{j+1,l-1}\\&=\sum_{k=1}^{M}\sum_{j+l=k}ljS_{jl}+\sum_{k=1}^{M}\sum_{j+l=k}jS_{jl},\end{aligned}$$

thus,

$$\frac{\sum_{k=1}^{M}\sum_{j+l=k}\big[l(j+1)S_{j+1,l-1}-ljS_{jl}\big]}{\sum_{k=1}^{M}\sum_{j+l=k}jS_{jl}}=1.$$

Hence,

$$\begin{aligned}\sum_{k=1}^{M}\sum_{j+l=k}lS'_{jl}&=-\sum_{k=1}^{M}\sum_{j+l=k}\beta l^2S_{jl}+\sum_{k=1}^{M}\sum_{j+l=k}\gamma lI_{jl}\\&\quad-\sum_{k=1}^{M}\sum_{j+l=k}\gamma lS_{jl}+\sum_{k=1}^{M}\sum_{j+l=k}j\beta lS_{jl}.\end{aligned}$$

Using a similar technique on (8b) gives

$$\frac{\sum_{k=1}^{M}\sum_{j+l=k}\big[j(j+1)I_{j+1,l-1}-j^2I_{jl}\big]}{\sum_{k=1}^{M}\sum_{j+l=k}jI_{jl}}=-1.$$

Hence,

$$\begin{aligned}\sum_{k=1}^{M}\sum_{j+l=k}jI'_{jl}&=\sum_{k=1}^{M}\sum_{j+l=k}j\beta lS_{jl}-\sum_{k=1}^{M}\sum_{j+l=k}\gamma jI_{jl}\\&\quad+\sum_{k=1}^{M}\sum_{j+l=k}\gamma lI_{jl}-\sum_{k=1}^{M}\sum_{j+l=k}\beta l^2S_{jl}.\end{aligned}$$

Thus,

$$\left[\sum_{k=1}^{M}\sum_{j+l=k} jI_{jl} - \sum_{k=1}^{M}\sum_{j+l=k} lS_{jl}\right]' = -\gamma\left[\sum_{k=1}^{M}\sum_{j+l=k} jI_{jl} - \sum_{k=1}^{M}\sum_{j+l=k} lS_{jl}\right].$$

Thus, if the initial conditions satisfy the balance condition (which must be so for a disease network), then the balance condition is satisfied for all further time.

Appendix C: Comparison of $\mathcal{R}_0$ for the SIS effective degree model and percolation theory

Consider a reduced SIS effective degree model in which the term $\beta(s+i)I_{s+1,i-1}$ in (8b) is set to zero. In this case the equations for I_{k0} uncouple from the other equations and so the order of the system is reduced to $\sum_{k=1}^{M} 2k = M(M+1)$. Now reorder the variables as in S_{si}, I_{si} pairs, i.e., $S_{01}, I_{01}; S_{11}, I_{11}, S_{02}, I_{02}; \ldots, I_{0M}$. Matrix V_k is now block upper triangular with diagonal blocks of order 2. Because of the sparsity of F, only the blocks on variables $S_{k-1,1}$ and $I_{k-1,1}$ contribute to FV^{-1}. These blocks of V are

$$\begin{bmatrix} \beta+\gamma & -\gamma \\ -\beta & 2\gamma+(k-1)\beta \end{bmatrix}$$

with inverse

$$\frac{1}{2\gamma^2+(k-1)\beta^2+k\beta\gamma}\begin{bmatrix} 2\gamma+(k-1)\beta & \gamma \\ \beta & \beta+\gamma \end{bmatrix}.$$

Thus for our reduced model $\mathcal{R}_0 = \rho(FV^{-1})$ gives

$$\mathcal{R}_0 = \sum_{k=1}^{M} \frac{2\beta\gamma+(k-1)\beta^2}{2\gamma^2+(k-1)\beta^2+k\beta\gamma}\frac{k(k-1)S_{k0}}{\sum_{k=1}^{M} kS_{k0}} > \frac{\beta}{\beta+\gamma}\frac{\langle k(k-1)\rangle}{\langle k\rangle}.$$

Hence, for the same parameters our reduced model $\mathcal{R}_0$ is greater than the percolation $\mathcal{R}_0$ given by (1).

To compare our reduced model with the SIS effective degree model given by (8), note that each entry of V_k for the reduced model is greater than or equal to the corresponding V_k entry for the effective degree model. By the theory of M-matrices, the inequality on corresponding entries of their inverse matrices are reversed. Since FV^{-1} is non-negative, $\mathcal{R}_0$ for the reduced model is less than or equal to $\mathcal{R}_0$ for the effective degree model. This proves that for the same parameters β, γ, and N_k, the $\mathcal{R}_0$ of the effective degree model (12) is always greater than that of the percolation theory (1).

References

Ball F (1983) The threshold behavior of epidemic models. J Appl Probab 20:227–241
Ball F, Neal P (2008) Network epidemic models with two levels of mixing. Math Biosci 212:69–87
Bansal S, Grenfell BT, Meyers LA (2007) When individual behaviour matters: homogeneous and network models in epidemiology. J R Soc Interface 4:879–891
Brauer F (2008) Compartmental models in epidemiology. In: Brauer F, van den Driessche P, Wu J (eds) Mathematical epidemiology, chapter 2. Springer, Berlin
Diekmann O, Heesterbeek JAP (2000) Mathematical epidemiology of infectious diseases: model building, analysis, and interpretation. Wiley, New York
Erdős P, Rényi A (1959) On random graphs. I. Publ. Math. 6:290–297
Fine PEM, Clarkson JA (1982) Measles in England and Wales—I: an analysis of factors underlying seasonal patterns. Int J Epidemiol 11:5–14
Gillespie DT (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J Comput Phys 22(4):403–434
Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. J Phys Chem 81(25):2340–2361
Kiss IZ, Green DM, Kao RR (2006) The effect of contact heterogeneity and multiple routes of transmission on final epidemic size. Math Biosci 203:124–136
Lloyd AL, Valeika S (2007) Network models in epidemiology: an overview. In: Blasius B, Kurths J, Stone L (eds) Complex population dynamics: nonlinear modeling in ecology, epidemiology and genetics, chapter 1. World Scientific, Singapore
Neal P (2007) Coupling of two SIR epidemic models with variable susceptibility and infectivity. J Appl Probab 44:41–57
Newman MEJ (2002) Spread of epidemic disease on networks. Phys Rev E 66:016128
Pastor-Satorras R, Vespignani A (2001a) Epidemic dynamics and endemic states in complex networks. Phys Rev E 63:066117
Pastor-Satorras R, Vespignani A (2001b) Epidemic spreading in scale-free networks. Phys Rev Lett 86: 3200–3203
Pastor-Satorras R, Vespignani A (2002) Epidemic dynamics in finite size scale-free networks. Phys Rev E 65:035108R
van den Driessche P, Watmough J (2002) Reproduction numbers and subthreshold endemic equilibria of compartmental models for disease transmission. Math Biosci 180:29–48
Varga RS (1962) Matrix iterative analysis. Prentice-Hall, Englewood Cliffs
Volz E (2008) SIR dynamics in random networks with heterogeneous connectivity. J Math Biol 56:293–310