J. Math. Biol.
DOI 10.1007/s00285-015-0962-4

Mathematical Biology

Disease invasion risk in a growing population

Sanling Yuan[1] · P. van den Driessche[2] · Frederick H. Willeboordse[3] · Zhisheng Shuai[4] · Junling Ma[2]

Received: 29 June 2015 / Revised: 6 December 2015

Abstract The spread of an infectious disease may depend on the population size. For simplicity, classic epidemic models assume homogeneous mixing, usually standard incidence or mass action. For standard incidence, the contact rate between any pair of individuals is inversely proportional to the population size, and so the basic reproduction number (and thus the initial exponential growth rate of the disease) is independent of the population size. For mass action, this contact rate remains constant, predicting that the basic reproduction number increases linearly with the population size, meaning that disease invasion is easiest when the population is largest. In this paper, we show that neither of these may be true on a slowly evolving contact network: the basic reproduction number of a short epidemic can reach its maximum while the population is still growing. The basic reproduction number is proportional to the spectral radius of a contact matrix, which is shown numerically to be well approximated by the average excess degree of the contact network. We base our analysis on modeling the dynamics of the average excess degree of a random contact network with constant population input, proportional deaths, and preferential attachment for contacts brought in by incoming individuals (i.e., individuals with more contacts attract more incoming contacts). In addition, we show that our result also holds for uniform attachment of incoming contacts (i.e., every individual has the same chance of attracting incoming contacts), and much more general population dynamics. Our results show that a dis-

✉ Junling Ma
junlingm@uvic.ca

1 College of Science, Shanghai University for Science and Technology, Shanghai, China

2 Department of Mathematics and Statistics, University of Victoria, Victoria, BC, Canada

3 University Scholars Programme, National University of Singapore, Singapore, Singapore

4 Department of Mathematics, University of Central Florida, Orlando, FL 32816, USA

Published online: 21 January 2016

ease spreading in a growing population may evade control if disease control planning is based on the basic reproduction number at maximum population size.

Keywords Dynamic contact network · Basic reproduction number · Excess degree

Mathematics Subject Classification 92D30

1 Introduction

Real populations are heterogeneous, some people have many friends and contacts and some people only have a few, and there are births, deaths and possibly migrations. However, detailed knowledge, let alone real-time knowledge, of this heterogeneity is hard if not impossible to aquire. Furthermore, even with such knowledge, a mathematical analysis using all the available information would be challenging to say the least. Therefore, classical disease models make a number of simplifications that, given the right circumstances, lead to very useful results.

One of the key simplifications of classical disease models is to assume a homogeneous population, where the probability of an individual to pair up with any other individual is the same, although this probability may change over time. With the mass action assumption, the contact rate per pair is constant, and thus the contact rate of an individual to the whole population increases linearly with the population size. An alternative assumption is standard incidence, which assumes that the contact rate of an individual to the whole population remains constant when the population changes. Thus the contact rate between any pair of individuals must decrease inversely proportional to the population size. See, e.g., Hethcote (2000) for a detailed discussion on these assumptions.

The basic reproduction number $\mathcal{R}_0$ represents the average number of secondary infections caused by a typical infectious individual introduced into a wholly susceptible population. In general, a disease can spread in a population if and only if $\mathcal{R}_0 > 1$, giving a disease threshold. The mass action assumption implies that $\mathcal{R}_0$, and thus the initial disease growth rate, increases linearly with population size; and the standard incidence assumption implies that $\mathcal{R}_0$ is constant when the population size changes. Classical models seldom predict that $\mathcal{R}_0$ is a non-monotonic function of the population size.

A more realistic model for contact heterogeneity is a contact network, which is a graph with nodes representing individuals and edges representing contacts between individuals. Homogeneous mixing is then represented by a complete graph. Thus classical disease models are special cases of network disease models. Even though the exact contact network is difficult to observe, network statistics, such as the degree (the number of neighbors of a node) distribution, are sometimes collectable; see, e.g., Pourbohloul et al. (2005). Networks that are represented only by their degree distribution can be represented by graphs generated by the configuration model (also called the Molloy and Read model, see Molloy and Reed 1995): Given a number of individuals (nodes) and a degree distribution P_k (the probability that a random node has degree k), for each node, draw an integer degree k from the distribution

P_k, and assign k "stubs" to the node; then uniformly randomly pick two "stubs" without replacement and connect them to form an edge; repeat this edge formation until no such edges can be formed. To avoid self-loops and multiple edges, drop such edges from the generated network. A network generated by this process typically has negligible clustering (triangles) and negligible degree correlation (correlation between the degrees of the nodes on the opposite ends of a random edge).

Predictions from network models may differ from those of classical models, for example, SIS and SIR effective degree network models have different disease thresholds due to different $\mathcal{R}_0$ expressions (Lindquist et al. 2010), and a scale-free network (one on which $P_k \propto k^{-\alpha}$ for $\alpha > 0$) has no disease threshold (Pastor-Satorras and Vespignani 2002; Chatterjee and Durrett 2009). For SIR disease dynamics on a static random network generated by a configuration model, $\mathcal{R}_0$ is equal to the probability of disease transmission along a random edge (T) times the average excess degree ($\langle k_e \rangle$), i.e., the degree of a node found by following a random edge less one (the edge that is followed); see, eg., Newman (2002a), Volz (2008), Miller (2011), Lindquist et al. (2010). That is,

$$\mathcal{R}_0 = T \langle k_e \rangle. \tag{1}$$

To understand this, consider a newly infected node that was infected by one of its neighbors (i.e., the node is found by following an edge from an infected neighbor). Because it cannot infect the neighbor who infected it, the number of neighbors that this newly infected node can infect is its degree minus one; see the mathematical definition of $\langle k_e \rangle$ for a random contact network in (7) in Sect. 2.

We consider a short epidemic on a slowly evolving network, so that the network can be considered static for the duration of the epidemic. We assume that on the network the contact rate along edges remains constant, analogously to mass action. When the population grows, new nodes are added to the population, and new edges are formed between the incoming nodes and the original nodes. When nodes leave the population, they also take away their edges. In fact, since each incoming node increases the total degree in the network by twice its degree, and each leaving node decreases the total degree also by twice its degree, for a growing network that approaches equilibrium, the total degree approaches equilibrium twice as fast as the number of nodes does. Thus, it is possible that the average degree in the network reaches a maximum then decreases. In this paper, we investigate the possibility that, on a contact network, the disease invasion risk as measured by the basic reproduction number reaches a maximum then decreases. Note that here we ignore the stochasticity of the disease spread process, i.e., ignore the probability that a disease may not invade the network due to stochasticity even if the basic reproduction number is greater than unity.

On a growing network, the disease dynamics are difficult to model as there are changes in degree distribution, and also degree correlation may be introduced. For simplicity, here we consider SIR type diseases such as a strain of influenza where an individual can be susceptible (S), infectious (I) or have recovered (R) with subsequent lifelong immunity. For some diseases (e.g., a strain of influenza, measles), birth, death and migration can usually be ignored during an epidemic due to the vast difference in time scales, since these diseases act on individuals on the time scale of days while birth

and death rates influence populations on the time scale of decades. For such diseases, the network can be regarded as static during an epidemic. Unfortunately, even on a static network, the basic reproduction number cannot be simply formulated as (1) because such a network typically has degree correlation; see, e.g., Britton et al. 2011, Section 3.3.3. Rather, we need to assume that the degrees i and j of nodes connected by a random edge have a joint distribution P_{ij}, so that

$$\mathcal{R}_0 = T\rho(C), \tag{2}$$

where ρ is the spectral radius operator, and C is the "contact matrix" with ij entry

$$C_{ij} = (j-1)\frac{P_{ij}}{\sum_k P_{ik}}. \tag{3}$$

This is equivalent to Equation (7) in Newman (2002b) in which the matrix A loses stability through a 0 eigenvalue. Ball et al. 2013, Equation (15) extend this contact matrix to include households. The value of $\mathcal{R}_0$ in (2) becomes much harder than (1) to evaluate even if the dynamics of P_{ij} can be modeled. If the degrees of the two nodes of a random edge are independent, then (1) and (2) are equal, since in this case the joint distribution matrix $P = [P_{ij}]$ has rank 1; see Appendix 2.

The dynamics of (2) on a slowly evolving network are difficult to study. For simplicity, in our analysis we use (1) as an approximation for the basic reproduction number on evolving networks. We formulate our model on the dynamics of the average excess degree in Sect. 2, and in Sects. 3 and 4 we discuss two scenarios on how new edges brought in by incoming nodes are attached to existing nodes. In particular, we prove that the basic reproduction number can indeed increase to a maximum, then decrease while the population increases. In Sect. 5, we show numerically that the approximation of $\rho(C)$ by $\langle k_e \rangle$ is good for these two scenarios. In Sect. 6, we show that our result on the basic reproduction number holds for more general population dynamics on networks. Concluding remarks are given in Sect. 7.

2 Model

Because the disease time scale is assumed to be much smaller than the network dynamics time scale, and (1) is used to approximate the basic reproduction number, the dynamics of $\mathcal{R}_0$ is only determined by the dynamics of the degree distribution of the network, specifically, the dynamics of the average excess degree $\langle k_e \rangle$.

To model the network dynamics, we classify the nodes by their degree. For $k \geq 0$, let N_k be the class of nodes with degree k, and let $N_k(t)$ be the number of nodes in this class at time t (for simplicity, we drop the time dependence). Here we develop a model for the dynamics of N_k based on our network rewiring model in Lindquist et al. (2009). This part of our model is similar to the population dynamics of the model in Jin et al. 2014, equation (15).

New individuals have a degree distribution π_k, and enter the population at rate λ, so that the rate at which they enter the N_k class is $\lambda\pi_k$. As they come in, they bring in

$\sum_k k\lambda\pi_k = \lambda\ell$ new edges per unit time, where $\ell = \sum_k k\pi_k$ is the average degree of the incoming nodes. Let $A_k(t)$ be the probability that a new edge is attached to any specific degree k node. Therefore, these new edges attach to nodes of N_k with an expected rate $\lambda\ell A_k N_k$, and cause these nodes to leave N_k and enter N_{k+1}. Similarly, N_{k-1} nodes have an edge attached and thus move to N_k with an expected rate $\lambda\ell A_{k-1}N_{k-1}$.

The death rate of each individual is μ, so that the rate at which nodes die in class N_k is μN_k. Due to the death of neighboring nodes, the number of $k+1$ degree nodes is reduced and these nodes become k degree nodes at rate $\mu(k+1)N_{k+1}$. For the same reason, k degree nodes become $k-1$ degree nodes at rate $\mu k N_k$.

Putting the above together with the assumption $N_{-1} = A_{-1} = 0$ gives for $k \geq 0$:

$$\frac{d}{dt}N_k = \lambda\pi_k + \lambda\ell(N_{k-1}A_{k-1} - N_k A_k) - \mu N_k + \mu[(k+1)N_{k+1} - kN_k]\,. \quad (4)$$

The total population is $N = \sum_k N_k$, which satisfies

$$\frac{d}{dt}N = \sum_k \frac{d}{dt}N_k = \lambda - \mu N\,. \quad (5)$$

The total degree (twice the total number of edges) is $L = \sum_k kN_k$, which satisfies

$$\frac{d}{dt}L = \sum_k k\frac{d}{dt}N_k = 2\lambda\ell - 2\mu L\,. \quad (6)$$

The average excess degree is defined as (see, e.g., Newman 2002a)

$$\langle k_e\rangle = \frac{\sum_k (k-1)kN_k}{\sum_k kN_k} = \frac{\sum_k (k-1)kN_k}{L}\,. \quad (7)$$

Note that $\langle k_e\rangle$ does not depend on the degree correlation on a random network, because the probability that a random edge leads to a degree k node (and thus has an excess degree $k-1$) is kN_k/L. A formal proof is given in Appendix 1 with a model that properly accounts for degree correlation.

Our goal is to model the dynamics of $\langle k_e\rangle$. Differentiate (7) with respect to time, and substitute in (4) and (6),

$$\begin{aligned}\frac{d}{dt}\langle k_e\rangle =& \frac{\sum_j j(j-1)\frac{d}{dt}N_j}{L} - \frac{\frac{d}{dt}L\sum_j j(j-1)N_j}{L^2}\\ =& \frac{1}{L}\sum_j j(j-1)\lambda\pi_j + \frac{\mu}{L}\sum_j j(j-1)(j+1)(N_{j+1}-N_j)\\ &+ \frac{\lambda\ell}{L}\sum_j j(j-1)(N_{j-1}A_{j-1} - N_j A_j) - \frac{2\lambda\ell - 2\mu L}{L}\langle k_e\rangle\\ =& \frac{\lambda\ell}{L}(E_\lambda + 2\langle k_a\rangle) - \left(\frac{2\lambda\ell}{L} + \mu\right)\langle k_e\rangle\,, \end{aligned} \quad (8)$$

where

$$E_\lambda = \frac{\sum_k k(k-1)\pi_k}{\ell}$$

is the average excess degree of an incoming node, and $\langle k_a \rangle = \sum_k k A_k N_k$ is the average degree of a randomly selected node for the attachment of a new edge.

For (6) and (8) to be meaningful, we assume ℓ, E_λ, and the initial conditions $L(0)$ and $\langle k_e \rangle$ (0) to be finite. To close our model, the probabilities A_k need to be defined. Here we consider two special cases, preferential attachment and uniform attachment, and we discuss them separately in the following two sections.

3 Preferential attachment

Preferential attachment assumes that a new edge is attached to an existing node with a probability proportional to its degree (see, e.g., Albert and Barabási 2002); in other words, famous people become famous quicker. Thus,

$$A_k = \frac{k}{L}. \tag{9}$$

Then, (4) can be written as

$$\frac{d}{dt}N_k = \lambda\pi_k + \frac{\lambda\ell}{L}[(k-1)N_{k-1} - kN_k] + \mu(k+1)(N_{k+1} - N_k). \tag{10}$$

In addition,

$$\langle k_a \rangle = \frac{\sum_k k^2 N_k}{L} = \langle k_e \rangle + 1.$$

Thus, (8) becomes

$$\frac{d}{dt}\langle k_e \rangle = \frac{\lambda\ell}{L}(E_\lambda + 2) - \mu \langle k_e \rangle. \tag{11}$$

The system (6) and (11) has a unique equilibrium that satisfies

$$\begin{aligned} 2\lambda\ell - 2\mu L^* &= 0, \\ \frac{\lambda\ell}{L^*}(E_\lambda + 2) - \mu \langle k_e \rangle^* &= 0, \end{aligned}$$

that is, $L^* = \lambda\ell/\mu = N^*\ell$ where $N^* = \lambda/\mu$ is the equilibrium population size, and $\langle k_e \rangle^* = E_\lambda + 2$.

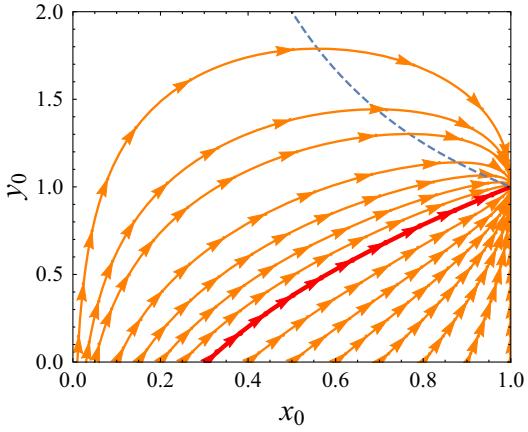

Fig. 1 Phase plot of the solutions of (12) for different initial conditions $x_0, y_0 \in (0, 1)$. The y value is proportional to the basic reproduction number $\mathcal{R}_0$. The maximum average excess degree $\langle k_e \rangle$ occurs on the *dashed curve* $xy = 1$. Above the *bold solution curve* the initial conditions satisfy (13)

To analyze the dynamics of (6) and (11), non-dimensionalize the system with the transformation $\tau = \mu t$, $L = L^* x$ and $\langle k_e \rangle = \langle k_e \rangle^* y$, giving

$$\frac{dx}{d\tau} = 2 - 2x, \tag{12a}$$

$$\frac{dy}{d\tau} = \frac{1}{x} - y. \tag{12b}$$

The phase plot of this simple system is illustrated in Fig. 1 for different initial conditions.

This system is solvable analytically. With the initial condition $x(0) = x_0$, (12a) gives

$$x(\tau) = 1 + (x_0 - 1)e^{-2\tau},$$

which can then be substituted in (12b), and with $y(0) = y_0$,

$$y(\tau) = 1 + (y_0 - 1)e^{-\tau} + \frac{1}{2}\sqrt{1 - x_0}\, e^{-\tau} \log \frac{(e^{\tau} - \sqrt{1 - x_0})(1 + \sqrt{1 - x_0})}{(e^{\tau} + \sqrt{1 - x_0})(1 - \sqrt{1 - x_0})}.$$

Note that, with $x_0 < 1$, $x(\tau)$ monotonically increases to $x = 1$. On the other hand, $\frac{dy}{d\tau}$ is positive for $1 < xy$ and negative for $1 > xy$. Thus, with $y_0 < 1$, $y(\tau)$ may either increase monotonically and reach a maximum on $xy = 1$ then decrease to $y = 1$, or increase monotonically to $y = 1$. Consequently, the behavior of $y(\tau)$ is uniquely determined by the sign of $\frac{dy}{d\tau}$ near the equilibrium. As $\tau \to \infty$,

$$\frac{dy}{d\tau} = -\left[(y_0 - 1) + \frac{1}{2}\sqrt{1 - x_0} \log \frac{1 + \sqrt{1 - x_0}}{1 - \sqrt{1 - x_0}}\right] e^{-\tau} + o(e^{-\tau}).$$

Hence, if

$$\frac{\sqrt{1-x_0}}{2}\log\frac{1+\sqrt{1-x_0}}{1-\sqrt{1-x_0}} > 1-y_0, \tag{13}$$

then $y(\tau)$ reaches its equilibrium $y=1$ from above ($\frac{dy}{d\tau}<0$ at $y=1$). Thus $y(\tau)$ reaches its maximum with $y>1$ and then decreases to $y=1$. If the above inequality is reversed, then $y(\tau)$ increases monotonically to $y=1$.

Note that the equilibrium degree distribution N_k^* could be determined from (10) at equilibrium with $L=L^*$, which is a linear system of equations with a tridiagonal coefficient matrix.

4 Uniform attachment

Uniform attachment assumes that the new edges brought in by the incoming nodes are uniformly attached to a node independent of its degree, i.e.,

$$A_k = \frac{1}{N}.$$

Then

$$\frac{d}{dt}N_k = \lambda\pi_k + \mu(k+1)[N_{k+1}-N_k] + \frac{\lambda\ell}{N}[N_{k-1}-N_k]. \tag{14}$$

Thus, (8) becomes

$$\frac{d}{dt}\langle k_e\rangle = \frac{\lambda\ell}{L}(E_\lambda + 2\langle k\rangle) - \left(\frac{2\lambda\ell}{L}+\mu\right)\langle k_e\rangle. \tag{15}$$

Note that the dynamics of $\langle k_e\rangle$ depends on that of $\langle k\rangle$, which is determined by

$$\frac{d}{dt}\langle k\rangle = \frac{d}{dt}\left(\frac{L}{N}\right) = \left(\frac{2\lambda\ell}{L}-\mu\right)\langle k\rangle - \frac{\lambda}{L}\langle k\rangle^2. \tag{16}$$

The system (6), (15), and (16) determines the dynamics of $\langle k_e\rangle$. This system has a unique positive equilibrium $L^*=\lambda\ell/\mu$, $\langle k\rangle^*=\ell$, and $\langle k_e\rangle^*=(E_\lambda+2\ell)/3$. Let $L=L^*x$, $\langle k_e\rangle=\langle k_e\rangle^* y$, $\langle k\rangle=\langle k\rangle^* z$, and $\tau=\mu t$. The non-dimensionalized system is

$$\begin{aligned}
\frac{dx}{d\tau} &= 2-2x,\\
\frac{dy}{d\tau} &= \frac{E_\lambda+2\ell z}{x\langle k_e\rangle^*} - \left(\frac{2}{x}+1\right)y,\\
\frac{dz}{d\tau} &= \left(\frac{2}{x}-1\right)z - \frac{z^2}{x}.
\end{aligned}$$

This system is also solvable with initial conditions $x(0) = x_0$, $y(0) = y_0$, $z(0) = z_0$, giving

$$
\begin{aligned}
x(\tau) &= 1 + (x_0 - 1)e^{-2\tau}, \\
z(\tau) &= \frac{1 + (x_0 - 1)e^{-2\tau}}{1 + (x_0/z_0 - 1)e^{-\tau}}, \\
y(\tau) &= \Big[-6\ell x_0 (x_0^2 + x_0 z_0^2 - 3x_0 z_0 - z_0^3 - 2z_0^2) \log \frac{e^\tau + x_0/z_0 - 1}{x_0/z_0} \\
&\quad + 6\ell x_0 z_0 (x_0 + z_0^2 - 2z_0)(e^\tau - 1) + 3\ell z_0^2 (z_0 - x_0)(e^{2\tau} - 1) \\
&\quad + (E_\lambda + 2\ell) z_0^3 (e^{3\tau} - 1) + 3x_0 y_0 z_0^3 \langle k_e \rangle^* \Big] \frac{1}{3z_0^3 \langle k_e \rangle^* [e^{3\tau} + (x_0 - 1)e^\tau]}.
\end{aligned}
$$

Expand $\frac{dy}{d\tau}(\tau)$ as a Taylor polynomial in $e^{-\tau}$,

$$
\frac{dy}{d\tau}(\tau) = \frac{\ell(x_0/z_0 - 1)}{\langle k_e \rangle^*} e^{-\tau} + O(e^{-2\tau}).
$$

Note that $x_0 = L_0/L^*$ and $z_0 = \langle k_0 \rangle / \ell = L_0/(N_0 \ell)$, so that $x_0/z_0 = N_0/N^*$ where $N^* = \lambda/\mu$ is the equilibrium population size. Thus, $x_0/z_0 \leq 1$ for a growing network (i.e., $N_0 < N^*$). This implies that $\frac{dy}{d\tau}(\tau) < 0$ as the orbits of the system approach the equilibrium, i.e., $y(t)$ decreases to equilibrium. Hence, in the case of a growing network, the average excess degree (and thus $\mathcal{R}_0$) always increases to a maximum then decreases to the equilibrium.

5 Numerical simulations

In this section, we first show numerically that (1) is a good approximation to (2). Note that a model for the joint distribution of the degrees of the nodes connected by a random edge, P_{ij}, on a growing random network is developed as (19) in Appendix 1, and is shown numerically to precisely capture the average dynamics of the underlying stochastic process (see Fig. 4). We thus solve this model numerically, and use the resulting $P_{ij}(t)$ to compute $\mathcal{R}_0$ using (2) and (3). This basic reproduction number is then compared numerically to the approximation (1), with the relation

$$
\langle k_e \rangle = \sum_{i,j} (i - 1) P_{ij},
$$

derived from (7), (20) and (21). Figure 2 shows such comparisons on an evolving network with a constant immigration (or birth) rate and a constant per-capita death rate. For both preferential attachment (Sect. 3) and uniform attachment (Sect. 4) of the new edges brought in by incoming nodes to existing nodes, the two values follow exactly the same trend. Even though Fig. 2 only shows the comparison of a single set of disease and population parameters, it represents the qualitative behavior of (1) and

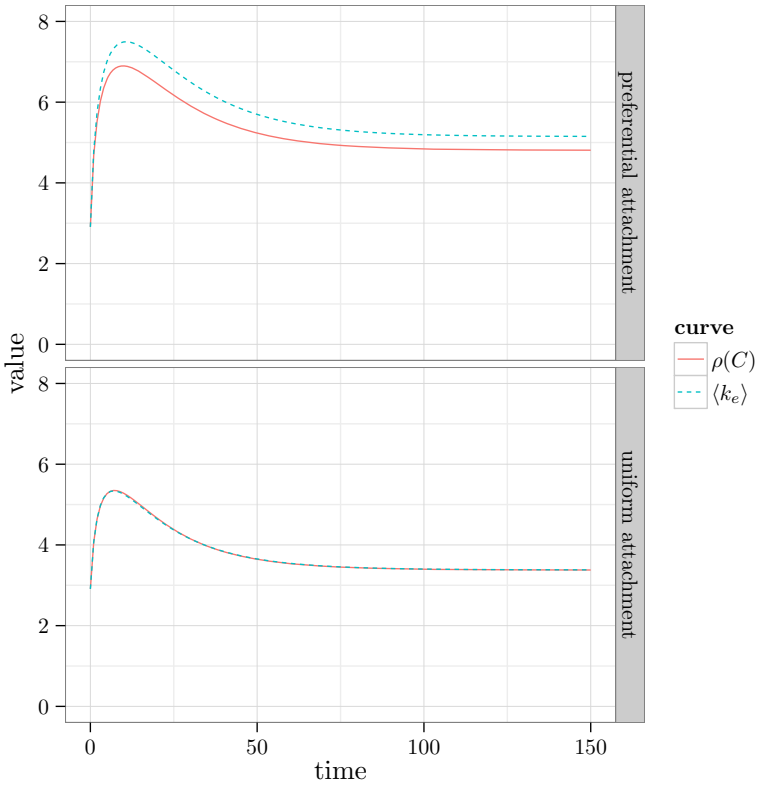

Fig. 2 The average excess degree $\langle k_e \rangle$ and the dominant eigenvalue of the contact matrix C as functions of time on an evolving network. The initial network has 2000 nodes and a Poisson degree distribution with a mean 3. The incoming nodes have degrees 2 or 5 with equal probabilities. The population has a constant immigration rate $\lambda = 1000$, and a constant per capita death rate $\mu = 0.05$

(2) across multiple sets of parameters that we tested. Thus the dynamics of (1) are a good approximation to the dynamics of (2).

In fact, Fig. 2 also shows that, for the set of parameters given, both $\langle k_e \rangle$ and $\rho(C)$ (and thus $\mathcal{R}_0$) increase to a maximum then decrease to an equilibrium, and this is true for both the preferential attachment and the uniform attachment. This indeed agrees with the predictions of our model in Sects. 3 and 4.

To show that the underlying stochastic disease spread process shows the same behavior, we conduct 100 runs of stochastic simulations on a growing random network with preferential attachment. We use the Gillespie method (Gillespie 1976, 1977) for the simulation, namely, the new nodes are assumed to arrive as a Poisson process with rate λ (and thus the waiting time for arrival are independently exponentially distributed with mean $1/\lambda$), each arriving node has a degree randomly drawn from a distribution $\{\pi_k\}$, and for each of the edges of the incoming node, it is randomly attached to an existing node with a probability proportional to the degree of the existing node. Each node has a lifetime independently exponentially distributed with a rate μ, and when the node dies, it takes away all its edges. To simulate the disease spread process, initially

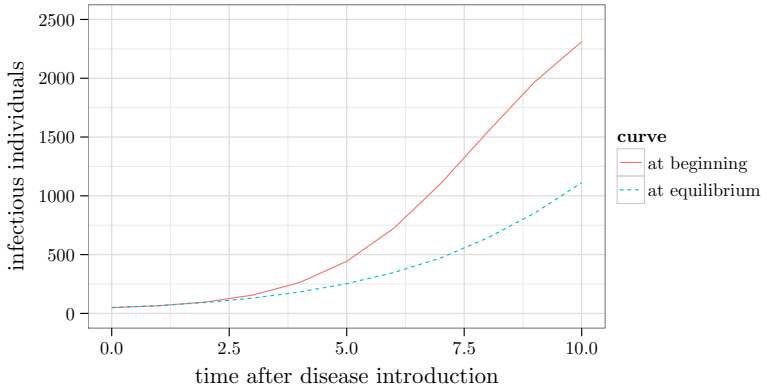

Fig. 3 Epidemic curves on a growing random contact network, for a disease introduced at the beginning of population growth (*solid curve*), and for a disease introduced at the population equilibrium (*dashed curve*). Time $t = 0$ corresponds to the time of disease introduction. Both curves are the average of 100 runs of stochastic simulations with the same disease parameters and contact network. The parameters of the network dynamics are the same as in Fig. 2 with preferential attachment. The transmission rate is $\beta = 0.15$, the infectious period is exponentially distributed with mean $1/\gamma = 5$, with $I_0 = 50$ initially infectious individuals

we introduce the disease by uniformly picking $I_0 = 50$ nodes and labeling them infectious while labeling the remaining nodes susceptible. The infected nodes contact each neighbor independently with a waiting time exponentially distributed with mean $1/\beta$, and upon contact, if the neighbor is susceptible, the neighbor is labeled infectious. Each infectious node, upon becoming infectious, is assigned an infectious period that is independently exponentially distributed with mean $1/\gamma$, at which time it is labeled recovered. There is a small probability that edge attachments lead to multiple edges or self loops. However, this probability approaches 0 as the network size approaches ∞. Thus, multiple edges can be ignored.

We consider two scenarios, in one the disease is introduced at the beginning of the growth for the network, and in the other the disease is introduced when the growth of the network reaches its equilibrium. Figure 3 shows the results of the ensemble average of 100 simulations for each scenario. It can be seen that a disease introduced during the early population growth phase spreads faster than a disease introduced at the population equilibrium with the same parameters. Note that, if the network dynamics are ignored, then the exponential growth rate is

$$r = \beta \langle k_e \rangle - \beta - \gamma,$$

which is a linear function of $\langle k_e \rangle$. This growth rate can be derived from the Volz-Miller model (Miller 2011).

6 More general demographics

With more general demographics for a population that stabilizes, assume that the number of new individuals coming into the population per unit time $b(N)$ and the

death rate of an individual $d(N)$ are independent of their degrees. The dynamics of the population size then becomes

$$\frac{d}{dt}N = b(N) - d(N)N. \tag{17}$$

In the special case that $b(N)$ and $d(N)$ are constants, i.e., $b(N) = \lambda$ and $d(N) = \mu$, this reduces to (5).

With the assumptions that an incoming node has degree k with probability π_k, and its k edges are randomly attached to a degree i node with probability A_i, the dynamics for N_k have the same form as (4), with λ replaced by $b(N)$ and μ replaced by $d(N)$. Thus

$$\frac{d}{dt}L = 2b(N)\ell - 2d(N)L.$$

Assume that $N(t) \to N^*$ as $t \to \infty$. Let $\lambda = b(N^*)$ and $\mu = d(N^*)$, then $L \to N^*\ell = \lambda\ell/\mu$. Linearizing $b(N)$ and $d(N)$ about N^*, the dynamics for N_k can be written as

$$\frac{d}{dt}N_k = F + O(N - N^*) \tag{18}$$

where F is the right hand side of (4). Thus, our model (4) is an approximation to the dynamics of (18) near the population equilibrium N^*, i.e., the qualitative behaviors of the two models are the same near $N = N^*$. Therefore, the behavior near L^* and $\langle k_e \rangle^*$ is similar to that of (6) and (8), implying that $\langle k_e \rangle$ (and thus $\mathcal{R}_0$) may increase to a maximum then decrease to the equilibrium value.

7 Discussion

We show that a growing population modeled by a random contact network can have a higher risk of disease invasion during the growth phase than at equilibrium and maximum population size, specifically, the basic reproduction number may be larger during the growth phase when the population size is smaller than the equilibrium size. This finding is in contrast to classical epidemic models where the basic reproduction number is largest at the maximum population size. Since contact networks are generally more realistic than the fully-linked networks of classical epidemic models, it can be expected that real world networks show this phenomenon as well. Thus in situations of rapid population growth, such as during a refugee crisis or a resource boom, comparatively more resources for disease management are necessary as compared to an established city even if the overall sanitation and other factors are similar.

Our approximation shows that the greater than classical (superlinear) increase of the basic reproduction number is caused by a rapid increase in the average excess degree of the growing contact network, since on a random contact network it is the average excess degree that determines the basic reproduction number of a disease with lifetime acquired immunity. This can be understood by considering that an infectious

node (except the few index cases) itself was infected by one of its neighbors through an edge, so that the number of secondary infections is determined by its degree less one (the edge to the neighbor who infected it). For any given disease, a large average excess degree implies a large basic reproduction number.

We develop a model for the time evolution of the average excess degree including recruitment and death, with the edges of incoming nodes preferentially attached to existing nodes (in other words, a new node links to an existing node with a probability that is proportional to the degree of the existing node). Our model shows that the average degree of the network evolves to the average degree of the incoming nodes (ℓ), and that the average excess degree evolves to the average excess degree (E_λ) of the incoming nodes plus two. If the initial average excess degree is smaller than the equilibrium value, the average excess degree on such a growing network can either increase monotonically to equilibrium, or first increase beyond equilibrium then decrease to equilibrium. The latter happens if the initial network has either a small number of edges or a large average excess degree. For such networks, the maximum average excess degree increases with the increase of the ratio between initial and equilibrium average excess degrees (y_0), and with the decrease of the ratio of the initial and equilibrium numbers of edges (x_0). Interestingly, this condition only depends on the initial values but not on the degree distribution of the incoming nodes. For more general demographics, as long as the population approaches a positive equilibrium, the above phenomenon is observed close to the equilibrium.

Our model verifies that recruitment and death can cause degree correlation (Fig. 4). Specifically, preferential attachment causes disassortative mixing, i.e., nodes with a small number of neighbors tend to connect to high degree nodes. For the recruitment and death taken in our model, the average excess degree is not affected by this correlation.

The average degree $\langle k \rangle = L/N$ can have a maximum during the growth phase because the number of edges grows twice as fast to equilibrium as the population size (see (5) and (6)), and thus has qualitatively the same behavior as the average excess degree. The dynamics of the average degree does not depend on the attachment scheme. This suggests that this behavior of $\langle k_e \rangle$ may happen for other attachment schemes for the edges of incoming nodes. We show that this is true for all growing networks in the case of uniform attachment.

In summary, we show that the maximum basic reproduction number can appear during the growth phase of a contact network, and furthermore show that this is not restricted to the assumed attachment scheme or population dynamics (as long as the population approaches a positive equilibrium). Compared to the classical notion that the disease invasion risk increases linearly with the population size, we find that the risk of disease invasion in a growing population manifests itself much earlier, i.e., when the population size is still relatively small. Consequently, response planning should make population growth an integral part of the factors considered.

Our model ignores the probability that the disease dies out before causing an epidemic even if the basic reproduction number $\mathcal{R}_0 > 1$. This effect is also an important factor of the disease invasion risk, and requires further investigation on a growing network.

Acknowledgments The research of JM and PvdD is partially funded by Canadian NSERC Discovery Grants. ZS acknowledges the support from a Simons Foundation Grant (#360178). SY is supported by the National Natural Science Foundation of China (11271260). The authors thank two anonymous reviewers for helpful comments that improved the exposition.

Appendix 1: Accounting for degree correlation

Stochastic simulations show that population dynamics may introduce degree correlation; see Fig. 4. Specifically, for preferential attachment, where the probability of attaching to a degree k node is proportional to k, the contacts become more disassortative (low degree nodes are more likely to connect with high degree nodes). To properly account for this degree covariance, we need to track the probability that a degree i node connects with a degree j node. For $i, j \geq 1$, let E_{ij} be the number of edges that connects two nodes with degrees i and j at time t.

The number of new edges brought in per unit time by new nodes with degree i is $\lambda i \pi_i$, and they attach to nodes with $j-1$ edges with probability $A_{j-1}N_{j-1}$, forming an edge in the E_{ij} class. At the same time, if a new edge attaches to a degree j node, the degree of the node becomes $j+1$, and thus E_{ij} edges become $E_{i,j+1}$ edges with rate $\lambda \ell A_i E_{ij}$. Similarly, $E_{i,j-1}$ edges become E_{ij} edges with rate $\lambda \ell A_{j-1} E_{i,j-1}$. Furthermore an E_{ij} edge can be formed if an edge of an incoming degree j node attaches to a node with degree $i-1$, while the existing $E_{i-1,m}$ edges of this degree $i-1$ node enter the E_{im} class.

Due to the death of either node of an E_{ij} edge, this edge disappears with rate 2μ. Alternatively, a neighbor of either of the nodes of an E_{ij} edge can die, with rate $(i-1)\mu$ and $(j-1)\mu$, respectively, and cause this E_{ij} edge to move to $E_{i-1,j}$ and $E_{i,j-1}$, respectively.

Putting the above together with the assumption $N_{-1} = A_{-1} = 0$ gives for $i, j \geq 1$

$$\begin{aligned}\frac{d}{dt}E_{ij} =& i\lambda\pi_i N_{j-1}A_{j-1} + j\lambda\pi_j N_{i-1}A_{i-1} + \mu\left[iE_{i+1,j} + jE_{i,j+1}\right.\\ &\left.-(i-1+j-1)E_{ij}\right] + \lambda\ell[A_{i-1}E_{i-1,j} + A_{j-1}E_{i,j-1} - (A_i + A_j)E_{ij}]\\ &- 2\mu E_{ij}. \end{aligned} \tag{19}$$

Initially $\sum_i E_{ij} = jN_j$ for all j, since the number of edges from a degree j node is jN_j, and the same quantity using E_{ij} notation is $\sum_i E_{ij}$. As a consistency check, we need to show that these two formulations are equivalent for all positive time. This can be verified by calculating

$$\begin{aligned}\frac{d}{dt}\sum_i E_{ij} =& \sum_i i\lambda\pi_i N_{j-1}A_{j-1} + j\lambda\pi_j \sum_i N_{i-1}A_{i-1}\\ &+ \mu\sum_i [iE_{i+1,j} + jE_{i,j+1} - (i+j)E_{ij}]\\ &+ \lambda\ell\sum_i [A_{i-1}E_{i-1,j} + A_{j-1}E_{i,j-1} - (A_i + A_j)E_{ij}]\end{aligned}$$

$$= \lambda j \pi_j + \mu \left[j \sum_i E_{i,j+1} - (j+1) \sum_i E_{ij} \right] + \lambda \ell \left[A_{j-1} \left(\sum_i E_{i,j-1} + N_{j-1} \right) - A_j \sum_i E_{ij} \right],$$

and with (4),

$$\begin{aligned} \frac{d}{dt} \left(\sum_i E_{ij} - jN_j \right) &= \mu j \left[\sum_i E_{i,j+1} - (j+1) N_{j+1} \right] \\ &\quad - \mu (j+1) \left(\sum_i E_{ij} - jN_j \right) \\ &\quad + \lambda \ell A_{j-1} \left[\sum_i E_{i,j-1} - (j-1) N_{j-1} \right] \\ &\quad - \lambda \ell A_j \left(\sum_i E_{ij} - jN_j \right). \end{aligned}$$

Thus, for each j, as it holds initially at time $t = 0$,

$$\sum_i E_{ij} = jN_j \tag{20}$$

holds for all time t.

Degree correlation

With E_{ij}, the joint distribution of the degrees of the two nodes connected by a random edge, P_{ij}, can be defined as

$$P_{ij} = \frac{E_{ij}}{\sum_{u,v} E_{uv}} = \frac{E_{ij}}{L} \tag{21}$$

where L is the total degree of the network. The degree correlation between a node of degree k_i and a node of degree k_j connected by a single edge is $\mathrm{Cor}[k_i, k_j] = \frac{\mathrm{Cov}[k_i, k_j]}{\sigma^2}$, where $\sigma^2 = \mathrm{Var}[k_i]$ is the variance of the excess degree distribution, and

$$\mathrm{Cov}[k_i, k_j] = \left\langle (k_i - \langle k_i \rangle)(k_j - \langle k_j \rangle) \right\rangle;$$

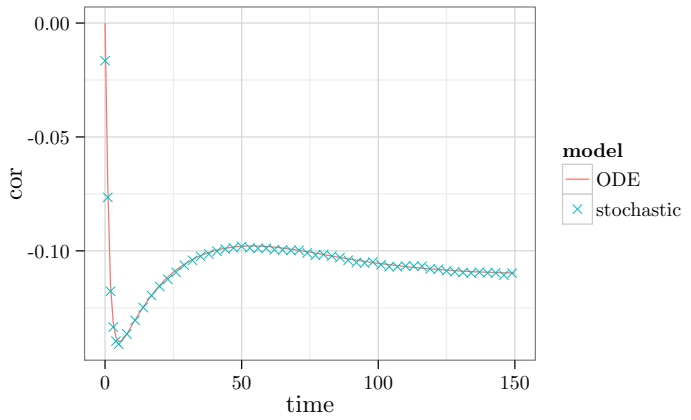

Fig. 4 Degree correlation caused by preferential attachment, comparison of 100 runs of stochastic simulations with the degree correlation formula given in (22). The network and model parameters are the same as in Fig. 3

see, e.g., (Newman 2002b). For our model this can be computed as

$$\mathrm{Cov}[k_i, k_j] = \sum_{i,j} ij P_{ij} - \left(\sum_{i,j} i P_{ij} \right)^2 . \tag{22}$$

Figure 4 shows that (22) agrees with stochastic simulations.

Average excess degree

The average degree of a neighbor is

$$\langle k_N \rangle = \frac{\sum_{i,j} j E_{ij}}{\sum_{i,j} E_{ij}} = \frac{\sum_j j (\sum_i E_{ij})}{\sum_j (\sum_i E_{ij})}.$$

Because $\sum_i E_{ij} = jN_j$,

$$\langle k_N \rangle = \frac{\sum_j j^2 N_j}{\sum_j jN_j} = \frac{\sum_j j^2 N_j}{L}.$$

Therefore, the average excess degree is

$$\langle k_e \rangle = \langle k_N \rangle - 1 = \frac{\sum_j j(j-1) N_j}{L}.$$

This shows that, even though the population dynamics may introduce degree correlation in E_{ij}, this correlation does not change the formula for the average excess degree.

Appendix 2: The basic reproduction number on networks with no assortativity

Under the assumption that the degrees of the two nodes of a random edge are independent, the joint distribution matrix of the two degrees $P = [P_{ij}] = P_i P_j$, where $P_i = \sum_j P_{ij} = iN_i/L$ is the probability that a random node found by following a random edge has degree i, for $i, j = 0, 1, \ldots, m-1$ where m is the largest degree of the network. Note that the matrix P has rank 1. Let the matrix $M = \text{diag}\{1/P_1, 1/P_2, \ldots, 1/P_m\}$, and $K = \text{diag}\{0, 1, 2, \ldots, m-1\}$, then the contact matrix $C = [C_{ij}]$ defined by (3) can be written as $C = MPK$, which also has rank 1. Thus, (2) becomes

$$\rho(C) = \rho(MPK) = \text{Trace}(MPK) = \sum_i (i-1)P_i = \sum_i (i-1)\frac{iN_i}{L} = \langle k_e \rangle .$$

That is, (1) and (2) are equivalent if the degrees of the two nodes of a random edge are independent.

References

Albert R, Barabási AL (2002) Statistical mechanics of complex networks. Rev Mod Phys 74:47–97

Ball F, Britton T, Sirl D (2013) A network with tunable clustering, degree correlation and degree distribution, and an epidemic thereon. J Math Biol 66:979–1019

Britton T, Lindholm M, Turova T (2011) A dynamic network in a dynamic population: asymptotic properties. J Appl Prob 48:1163–1178

Chatterjee S, Durrett R (2009) Contact processes on random graphs with power law degree distributions have critical value 0. Ann Prob 37:2332–2356

Gillespie DT (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J Comput Phys 22:403–434

Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. J Phys Chem 81:2340–2361

Hethcote HW (2000) The mathematics of infectious diseases. SIAM Rev 42:599–653

Jin Z, Sun G, Zhu H (2014) Epidemic models for complex networks with demographics. Math Biosci Eng 11:1295–1317

Lindquist J, Ma J, van den Driessche P, Willeboordse FH (2009) Network evolution by different rewiring schemes. Physica D 238:370–378

Lindquist J, Ma J, van den Driessche P, Willeboordse FH (2010) Effective degree network disease models. J Math Biol 62:143–164

Miller JC (2011) A note on a paper by Erik Volz: SIR dynamics in random networks. J Math Biol 62:349–358

Molloy M, Reed B (1995) A critical point for random graphs with a given degree sequence. Random Struct Algorithm 6:161

Newman MEJ (2002a) Spread of epidemic disease on networks. Phys Rev E 66:016128

Newman MEJ (2002b) Assortative mixing in networks. Phys Rev Lett 89:208701

Pastor-Satorras R, Vespignani A (2002) Epidemic dynamics in finite size scale-free networks. Phys Rev E 65:035108(R)

Pourbohloul B, Meyers LA, Skowronski DM, Krajden M, Patrick DM, Brunham RC (2005) Modeling control strategies of respiratory pathogens. Emerg Infect Dis 11:1249

Volz EM (2008) SIR dynamics in random networks with heterogeneous connectivity. J Math Biol 56:293–310