Electron. J. Probab. **18** (2013), no. 54, 1–30.
ISSN: 1083-6489 DOI: 10.1214/EJP.v18-2557

Approximating the epidemic curve

A. D. Barbour* G. Reinert†

Abstract

Many models of epidemic spread have a common qualitative structure. The numbers of infected individuals during the initial stages of an epidemic can be well approximated by a branching process, after which the proportion of individuals that are susceptible follows a more or less deterministic course. In this paper, we show that both of these features are consequences of assuming a locally branching structure in the models, and that the deterministic course can itself be determined from the distribution of the limiting random variable associated with the backward, susceptibility branching process. Examples considered include a stochastic version of the Kermack & McKendrick model, the Reed–Frost model, and the Volz configuration model.

Keywords: Epidemics; Reed–Frost; configuration model; deterministic approximation; branching processes.
AMS MSC 2010: Primary 92H30, Secondary 60K35; 60J85.
Submitted to EJP on January 14, 2013, final version accepted on May 2, 2013.

1 Introduction

Kermack & McKendrick's (1927) model of the course of an epidemic in a closed population has proved to be both effective in practice (see for example Brauer (2005), Brauer & Castillo–Chavéz (2012) p.350, Gupta *et al.* (2011)) and influential in the theoretical development of epidemic modelling. Writing $s(t)$ to denote the *density* of susceptible individuals in the population at time t and $\beta(v)$ the infectivity of an individual at time v after becoming infected, and normalizing the initial population density to be $s(-\infty) = 1$, the development of s is given by the equation

$$(-Ds(t)) \;=\; s(t)\int_0^\infty \beta(v)(-Ds(t-v))\,dv. \tag{1.1}$$

Here, Ds denotes the derivative of s with respect to time, and is *negative*. The quantity $(-Ds(t))$ is the rate at which the density of susceptibles is being reduced at time t, and this is just the (density standardized) rate at which infections are being made,

*Universität Zürich & National University of Singapore; E-mail: a.d.barbour@math.uzh.ch
†University of Oxford; E-mail: reinert@stats.ox.ac.uk

explaining the integral on the right hand side of (1.1) as the force of infection at time t. Dividing both sides of (1.1) by $s(t)$ and integrating gives

$$-\log s(t) \;=\; \int_0^\infty \beta(v)\{1-s(t-v)\}\,dv\,. \tag{1.2}$$

Note that, if s satisfies (1.2), so does any translate $\tilde{s}_h$ defined by $\tilde{s}_h(t) = s(t+h)$, for any $h \in \mathbb{R}$. However, it is shown in Diekmann (1977) that there is exactly one solution s to (1.2) that is non-increasing and non-negative, if, for instance, the value of $s(0) \in (0,1)$ is specified. Letting $t \to \infty$, (1.2) gives the final size equation

$$-\log s(\infty) \;=\; R_0(1-s(\infty))\,, \tag{1.3}$$

where the basic reproduction number $R_0 := \int_0^\infty \beta(v)\,dv$ is the expected total number of infections made by an infected individual in a susceptible population of unit density; a proportion $1-s(\infty)$ of the population has been infected by the end of the epidemic. Kermack & McKendrick (1927) then deduced their famous *threshold theorem*, that $s(\infty) < 1$ is only possible if $R_0 > 1$.

The final size equation can be interpreted more directly, without integrating (1.1), but at the level of an individual. Rewrite (1.3) in the form

$$s(\infty) \;=\; e^{-R_0(1-s(\infty))}, \tag{1.4}$$

and recognize $R_0(1-s(\infty))$ as the total integrated force of infection over the whole course of the epidemic. The tacit assumption about force of infection at the level of the individual is that it represents the 'instantaneous rate' of infection of an individual, interpreted in a Markovian sense, so that the probability of an individual avoiding infection after exposure to an integrated force of infection f should be given by e^{-f}. Thus the right hand side of (1.4) is the probability that an individual avoids infection throughout the whole course of the epidemic, which is exactly the proportion $s(\infty)$ that remain uninfected to the end.

The equation (1.4), with $s(\infty)$ replaced by the symbol q, also has a classic interpretation in a branching process context. It represents the equation for the extinction probability q of a branching process starting with a single individual, when the number of offspring has the Poisson distribution $\mathrm{Po}\,(R_0)$ with mean R_0. At an individual level, this suggests a stochastic analogue of the Kermack–McKendrick model, in which each infected individual makes potentially infectious contacts according to a Poisson process of rate $\beta(v)$, where v represents the time since infection. Each such event leads to a new infection, if the individual contacted is susceptible. In the early stages of an epidemic, almost all individuals are still susceptible, and so the early development of the epidemic is well approximated by a branching process, in which an individual at age v has (Markovian) birth rate $\beta(v)$. Branching processes have long been used to approximate the early stages of epidemic processes in this way. The earliest papers are those of Kendall (1956) and Whittle (1955), and a systematic treatment is given in Ball & Donnelly (1995). In particular, the Kermack–McKendrick threshold theorem is replaced by a stochastic threshold theorem, in which the probability that a large epidemic takes place, when started by a single infected individual K_0 in an initially susceptible population of large size N, is (approximately) $1-q$, thus being positive exactly when the mean number of offspring, here R_0, exceeds 1.

In contrast, for the analysis of the final size $N-S(\infty)$, where $S(t)$ denotes the number of susceptibles at time t, the appropriate branching approximation is not at the beginning of the epidemic, but approximates the process of the contacts potentially leading to the infection of a randomly chosen individual K — see, for example, Diekmann &

Heesterbeek (2000), pp. 171–172. If this backward process of contacts contains few individuals, as when its branching approximation dies out, then K is unlikely to become infected, whereas, if it contains many individuals, as when the branching approximation never dies out, K is almost certain to become infected, if the epidemic is a large one. Thus the probability that a randomly chosen K does not become infected is approximately 1 if the epidemic starting from K_0 is small, and approximately the extinction probability q_b for the 'backward' branching process, if the epidemic is large. However, because of the random choice of K, the probability that K escapes infection is just $N^{-1}\mathbb{E}S(\infty)$. Hence

$$1 - N^{-1}\mathbb{E}S(\infty) \;\approx\; 1 - \{q + (1-q)q_b\} \;=\; (1-q)(1-q_b),$$

so that, given that the epidemic starting from K_0 is a large one, the (mean of the) final proportion of infected individuals is close to $(1 - q_b)$. Thus, for $R_0 > 1$, a large epidemic is certain in the deterministic model, and the proportion of the population that is infected is $(1 - q_b)$. In the stochastic model, a large epidemic occurs only with probability approximately $(1 - q)$, in which case a proportion of approximately $(1 - q_b)$ of the individuals are infected, and on the complementary event there is only a tiny outbreak involving a negligible proportion of infected individuals. However, if the epidemic were started with $I > 1$ individuals, the probability of a large outbreak, again leading to a proportion of approximately $(1 - q_b)$ of the individuals being infected, increases to $(1 - q^I)$, and is thus nearly a certain event if I is at all large.

As it happens, for the particular stochastic Kermack–McKendrick model described above, the forward and backward branching processes are the same, so that $q_b = q$, and (1.4) is still the relevant equation for determining the final outcome of the epidemic, with $s(\infty)$ replaced by q_b. For most stochastic epidemic models, including the Markovian SIR model, the backward and forward branching processes are different.

In this paper, we use analogous ideas to show that, under appropriate assumptions, the whole course of the stochastic epidemic is determined by the analysis of the two branching processes, forward and backward. There is an initial phase, approximated as usual by the forward branching process. If this branching process does not become extinct, it settles to an essentially deterministic course of exponential growth, after a random delay that results from the initial random development of the branching process. After the point at which the forward branching process ceases to be a good approximation, the proportion of susceptibles in the epidemic process follows an almost deterministic development, which can be expressed in terms of properties of the backward branching process. One of the consequences of this is to show that the Markovian stochastic interpretation of the instantaneous force of infection, which is implicit in the derivation of the deterministic Kermack–McKendrick equation (1.1), is not actually necessary to justify the equation; we prove that (1.1) holds as a faithful approximation in a much wider class of models.

As an introduction to our approach, we now illustrate its application to the Reed–Frost discrete generation epidemic model in a population of size N. In this model, at any given time step, each susceptible individual avoids infection from each of the infectives with probability $(1 - p)$, with independence between infection attempts; infective individuals are removed in the next generation. Hence if at time t there are X_t susceptibles and Y_t infectives in the population, then the number of susceptibles at time $t+1$ has the Binomial distribution $\mathrm{Bi}\,(X_t, (1-p)^{Y_t})$. The epidemic ends when there are no infected individuals left in the population. Let the probability of an infected individual infecting a given susceptible be $p = \mu/N$. Then the approximating Galton-Watson forward branching process has offspring distribution $\mathrm{Po}\,(\mu)$ (and $R_0 = \mu$); we take $\mu > 1$. After n time units, the number of individuals alive in the branching process is $Z_n \sim W\mu^n$ and

the total number of individuals that were alive in previous generations is approximately $W\mu^n/(\mu-1)$, where W is the a.s. limit of $Z_n\mu^{-n}$. Take

$$n \;=\; n(N) \;:=\; \lfloor \tfrac{1}{2}\log N/\log\mu \rfloor,$$

so that $\mu^n = \theta_N N^{1/2}$, where $1 \le \theta_N < \mu$, and suppose that $W > 0$. Label those that have died in chronological order, with labels drawn independently and at random from $[N] := \{1, 2, \ldots, N\}$. Mark any whose labels have been used before, and all of their descendants, as 'ghosts'. There are only few marked, and those that are unmarked are the individuals that have been infected before time n in the epidemic. Let the set of labels used be denoted by L_N; its size is small compared to N.

Now, starting from a randomly chosen individual, take an independent realization of the reversed branching process — in this model, it has the same law as the forward process — and run it for $n(N) + r$ generations, after which there have been approximately $\widehat{W}\mu^{n+r+1}/(\mu-1)$ individuals born in total, where $\widehat{W}$ is the corresponding realization of the limit random variable, and is independent of W. Label these individuals, in order of appearance in this branching process, with labels drawn, with replacement, at random from $[N] \setminus L_N$, and again mark the (few) ghosts; let the set of labels be L_N^b, and denote by K the label of the initial individual. Do the same for the individuals alive in generation n of the forward process, and call this set L_N^f. If $L_N^b \cap L_N^f \neq \emptyset$, and an element of the intersection is a non-ghost, we can construct a chain of infection to it from the initial individual in the epidemic, and a chain going from it to K, giving a chain of infection from the start of the epidemic to K. Conversely, any chain of infection from the start of the epidemic to K must pass through a non-ghost element of $L_N^b \cap L_N^f$. Thus there is no chain of infection from the start of the epidemic to K exactly when $L_N^b \cap L_N^f$ is empty or contains only ghosts; the event that $L_N^b \cap L_N^f$ is non-empty but contains only ghosts has only small probability.

Now, given Z_n and the realization of the backward branching process, the mean number of intersections between L_N^b and L_N^f is close to $N^{-1}Z_n\,\widehat{W}\mu^{n+r+1}/(\mu-1)$, and hence, using a Poisson approximation, the probability of the intersection being empty is close to

$$\exp\{-N^{-1}Z_n\,\widehat{W}\mu^{n+r+1}/(\mu-1)\} \quad = \quad \exp\{-N^{-1/2}Z_n\,\widehat{W}\theta_N\mu^{r+1}/(\mu-1)\}.$$

It is now easy to convert this result into the statement

$$\begin{aligned}&\mathbb{P}[K \text{ has escaped infection until generation } 2n+r \,|\, \mathcal{F}_n]\\ &\quad\sim\; \mathbb{E}\big\{\exp\{-N^{-1/2}Z_n\,\widehat{W}\theta_N\mu^{r+1}/(\mu-1)\} \,|\, \mathcal{F}_n\big\},\end{aligned}$$

where K is a randomly chosen label from all of $[N]$ and $\mathcal{F}_n$ denotes $\sigma(Z_l,\, 0 \le l \le n)$; in other words, still with $n = n(N)$,

$$\mathbb{E}\{N^{-1}S_N(2n+r) \,|\, \mathcal{F}_n\} \;\sim\; \psi(N^{-1/2}Z_n\theta_N\mu^{r+1}/(\mu-1)),$$

where $S_N(t)$ denotes the number of susceptibles in the epidemic at generation t and $\psi(\theta) := \mathbb{E}\{e^{-\theta\widehat{W}}\}$.

But now, for two independently randomly chosen individuals K and K',

$$\begin{aligned}&\mathbb{E}\{(N^{-1}S_N(2n+r))^2 \,|\, \mathcal{F}_n\}\\ &\qquad=\; \mathbb{P}[\text{both } K \text{ and } K' \text{ have escaped infection until generation } 2n+r \,|\, \mathcal{F}_n]\end{aligned}$$

can be approximated in exactly the same way; since there is little overlap between the labels assigned to the backward branching processes starting from K and K', it is easy to deduce that

$$\mathbb{E}\{(N^{-1}S_N(2n+r))^2 \,|\, \mathcal{F}_n\} \;\sim\; \{\psi(N^{-1/2}Z_n\theta_N\mu^{r+1}/(\mu-1))\}^2$$

also, implying that $\mathrm{Var}\{N^{-1}S_N(2n+r)\,|\,\mathcal{F}_n\} \to 0$. Writing $W = \lim_{m\to\infty} Z_m\mu^{-m}$, we note that $Z_n = Z_{n(N)} \sim W\mu^{n(N)} = \theta_N N^{\frac{1}{2}} W$; this implies that, for any $\varepsilon > 0$ and any $r \in \mathbb{Z}$,

$$\lim_{N\to\infty} \mathbb{P}[|N^{-1}S_N(2n(N)+r) - \psi(W\theta_N^2\mu^{r+1}/(\mu-1))| > \varepsilon] \;=\; 0. \tag{1.5}$$

The quantity $\psi(W\theta_N^2\mu^{r+1}/(\mu-1))$ is random only through the presence of W. By time $n(N)$ the quantity W is essentially determined, and is the same for all $r \in \mathbb{Z}$. If $W = 0$, the above approximation is by $\psi(0) = 1$ for all r, indicating that only a small epidemic occurs; the assumption $\mu > 1$ merely ensures that $\mathbb{P}[W > 0] > 0$, so that a large epidemic is indeed possible.

If $W > 0$, one could describe the approximation slightly differently. The values of $N^{-1}S_N(2n(N)+r)$ for $r \in \mathbb{Z}$ are then approximated by a discrete subset of points on the continuous deterministic curve $u \mapsto \psi(\mu^{u+1}/(\mu-1))$, namely those with u of the form $r + \{\log W + 2\log\theta_N\}/\log\mu$ for $r \in \mathbb{Z}$. Thus randomness appears only as a time shift in the lattice of integer spaced points along the continuous deterministic path that are used for the approximation to the discrete time process. Note also that the times l at which $N^{-1}S_N(l)$ is not close either to 0 or to 1 are within $O(1)$ of $\log N/\log\mu$; the development of the epidemic is slow until almost time $\log N/\log\mu$, and then runs its course over comparatively few time steps.

In what follows, we shall make these arguments precise, but for processes with non-lattice offspring distributions in continuous time. The phenomena associated with discretization disappear, giving a neater result, but connecting the forward and backward branching processes becomes more delicate. Our analogue of (1.5) is proved in Theorem 2.10, under some fairly mild assumptions on the individual point processes of infection that include the stochastic Kermack–McKendrick model described above for many choices of the infectivity function β. It establishes that

$$\lim_{N\to\infty} \mathbb{P}[\sup_u |N^{-1}S_N(\lambda^{-1}\{\log N - \log W + u\}) - \hat{s}(u)| > \varepsilon] \;=\; 0, \tag{1.6}$$

for a deterministic function $\hat{s}$, whenever $W > 0$; here, λ is the Malthusian parameter (assumed positive) and W the limiting random variable for the associated forward branching process, and $\hat{s}$ is determined by the properties of the associated backward branching process. The methods that we use have quite general application, and have already been exploited in Barbour & Reinert (2012) in the context of the Aldous (2010) gossip process and of the Moore & Newman (1999) small world model. Related results, for first–passage percolation on the configuration graph, have been obtained by Bhamidi, van der Hofstad & Hooghiemstra (2012), also using branching process approximations.

The key ingredients that make the proofs go through are the branching nature of the forward and backward processes, and their exponential growth and stability properties. These are also shared, for instance, by their multitype analogues. We give a multitype analogue of (1.6) in Section 3.1, and discuss a configuration model in Section 3.2.

2 The single type model

2.1 The branching processes

We begin by considering an epidemic in a closed population of N individuals, where the population size N is to be thought of as large, that evolves according to the following scheme. Each individual i, $1 \le i \le N$, is equipped with a potential infection history, in the form of a realization of a point process ξ_i on $(0,\infty)$. If i becomes infected at time $\sigma(i) < \infty$, it makes infectious contacts with other individuals at times $\sigma(i,j) := \sigma(i) + \tau(i,j)$, where the times of the events of ξ_i are denoted by $0 < \tau(i,1) \le \tau(i,2) \le$

$\cdots$, and their number by $\nu(i) := \xi_i(\mathbb{R}_+) < \infty$; if required, ξ_i can be augmented by a time $\tau^R(i) \geq \tau(i, \nu(i))$, indicating that i is removed from the infectious state at time $\sigma(i) + \tau^R(i)$. The individuals contacted are chosen independently at random from $[N]$, and an infectious contact only results in the individual contacted becoming infected if they have not previously been contacted. The epidemic begins with individual i_1 becoming infected at time $\sigma(i_1) = 0$. After the r-th individual i_r has become infected at time $\sigma(i_r)$, and if $r < N$, then potential infectious contacts occur at the times $\sigma(i_r) + v_r(j)$, $1 \leq j \leq |V_r|$, where the $v_r(j)$ are the elements of

$$V_r \ := \ \{\sigma(i_l, j) - \sigma(i_r),\, 1 \leq j \leq \nu(i_l),\, 1 \leq l \leq r\} \cap (0, \infty),$$

arranged in non-decreasing order, and the labels of the individuals to be contacted are given by $I_r(j)$, $j \geq 1$, chosen independently and uniformly on $[N]$. Defining the index $j_*(r) := \min\{1 \leq j \leq |V_r| \colon\ I_r(j) \notin \{i_1, i_2, \ldots, i_r\}\}$, then

$$i_{r+1} := I_r(j_*(r)) \quad \text{and} \quad \sigma(i_{r+1}) \ = \ \sigma(i_r) + v_r(j_*(r)),$$

unless there is no such index $j_*(r)$, in which case the epidemic stops. It is assumed that $(\xi_i,\, 1 \leq i \leq N)$ are independent and identically distributed.

If the labelling were ignored, and $j_*(r)$ were taken to be 1 for each $r \geq 1$, and if the r-th infected individual were assigned infection history ξ'_r, with the $(\xi'_r,\, r \geq 1)$ independent and identically distributed, then the resulting process would be a Crump–Mode–Jagers branching process Z. Indeed, if the ξ'_r are distributed in the same way as ξ_1, the paths of the branching and epidemic processes (neglecting the labelling) can be coupled so as to agree exactly until $\rho := \min\{r \geq 0;\, j_*(r) \geq 2\}$ (Ball 1983, Ball & Donnelly, 1995), with the epidemic process recoverable from the branching process by adding labelling, and by marking as 'ghosts' individuals infected in the branching process but not in the epidemic process — $(j_*(r) - 1)$ such infections occur whenever $j_*(r) \geq 2$ — together with the individuals in the branching process that are descended from such individuals. We shall make substantial use of this coupling, but only up to times where there have typically been relatively few ghosts created.

We shall make the following assumptions on the distribution of ξ_1 of the above Crump–Mode–Jagers branching process. Let $p_j := \mathbb{P}[\nu(1) = j]$ and $\mu = \mathbb{E}\nu(1)$; denote the relative intensity measure of ξ_1 by

$$G(dt) \ := \ \mu^{-1}\mathbb{E}\xi_1(dt). \tag{2.1}$$

Assumptions

1. We assume that the branching process is supercritical, and that

$$1 \ < \ \mu \ < \ \infty; \qquad m_2 \ := \ \mathbb{E}\nu(1)^2 \ < \ \infty.$$

Let $\lambda > 0$ denote the Malthusian parameter of the branching process, satisfying

$$\mathbb{E}\left(\int_0^\infty e^{-\lambda t}\, \xi_i(dt)\right) \ = \ 1. \tag{2.2}$$

The existence of $\lambda > 0$ follows from Jagers (1975), Theorem 6.3.3, pp.131–2. We write

$$m_* \ := \ \mu\lambda \int_0^\infty t e^{-\lambda t}\, G(dt) \ < \ \infty; \tag{2.3}$$

then m_*/λ represents the mean age at child bearing (Jagers (1989), p.195).

2. The relative intensity measure G is non-lattice and has finite second moment. The support of G is a finite or semi-infinite open interval (a, b), and $G(A) \geq \int_A g(x)\,dx$ for any $A \subset (a, b)$, for some continuous positive density g. If $b = \infty$, then also $g(x) \geq kx^{-\gamma}$ for all $x \geq x_0$, for some $x_0 > a$, $k > 0$ and $\gamma > 3$.

Remark 2.1. *The tail condition for $b = \infty$ in Assumption 2 is necessitated by the method of proof for Corollary 2.5, and is presumably unnecessary for the results to hold. Indeed, Theorem 2.10 can be proved for the Markovian SIR epidemic using a different approach, although its relative intensity measure $G(dt) = \mu e^{-\mu t}\,dt$ has $b = \infty$, and does not satisfy our tail condition. Our method is designed to allow the wide range of dependence structures within the point processes ξ that might occur in practice. Assuming a bounded infectious period seems not to be too high a price to pay for this, at least in practical terms. The fact that fat tails are also covered can be viewed as a bonus.*

Remark 2.2. *Strictly speaking, the epidemic might be better modelled by assuming that the labels assigned to the individuals infected by any given individual i are chosen at random* without replacement from the labels *excluding i, and indeed that the number infected by a single individual cannot exceed $N - 1$. However, under the assumption that $m_2 < \infty$, the total variation distance between this distribution of labels and that being assumed here is at most $\frac{1}{2}N^{-1}(m_2 + \mu)$. Since, as will become clear, we shall need only to consider the offspring of at most $N^{5/8}$ individuals in our calculations, any difference between the results of the two models occurs with probability of order at most $O(N^{-3/8})$, and does not affect the results proved in this paper.*

Letting the infection times in the branching process be denoted by $(\sigma'(r),\, r \geq 1)$, and writing

$$B'(t) \;:=\; \max\{r\colon\, \sigma'(r) \leq t\} \tag{2.4}$$

for the number of births that have occurred in the branching process by time t, it follows that $W(t) := B'(t)e^{-\lambda t} \to W$ a.s. for a non-negative random variable W, (Nerman (1981), Theorem 5.4), and also that $\{W > 0\} = \{\lim_{t\to\infty} B'(t) = \infty\}$ a.s. (see (3.10) in Nerman (1981)). From Corollary 5.6 in Nerman (1981), and the fact that pointwise convergence to a continuous limit of non-decreasing bounded functions on $[0, \infty]$ is always uniform (Jagers (1975), p.170), it also follows that the statistics of the set

$$V'(t) \;:=\; \{\sigma'(l) + \tau'(l, j) - t,\, 1 \leq j \leq \nu'(l),\, 1 \leq l \leq B'(t)\} \cap (0, \infty),$$

where $\tau'(l, j)$ denotes the j-th point of ξ'_l and $\nu'(l) := \xi'_l(\mathbb{R}_+)$, converge in distribution, as $t \to \infty$, in the sense that, on $\{W > 0\}$,

$$\lim_{t\to\infty} \sup_{s\geq 0} \big|\, (|V'(t) \cap (0, s]|/|V'(t)|) - F(s) \big| \;=\; 0 \quad \text{a.s.}\,. \tag{2.5}$$

The set $V'(t)$ contains the times until birth of the unborn offspring of individuals born before t, and F is the distribution function on $\mathbb{R}_+$ given by

$$1 - F(s) \;:=\; \frac{\mu}{\mu - 1} \int_s^\infty (1 - e^{-\lambda(u-s)})\, G(du); \tag{2.6}$$

we have used Nerman (1981, Theorem 6.3) with $\phi := \chi_s$ and $\psi := \chi_0$, where $\chi_s(t) := \xi(s + t, \infty)\mathbf{1}\{t \geq 0\}$.

For the epidemic, the corresponding quantities depend on the choice of N, because of the role played by the labelling in its definition. We define

$$B_N(t) \;:=\; \max\{r\colon\, \sigma(i_r) \le t\}$$

and, in the natural notation,

$$V_N(t) \;:=\; V_{B_N(t)} + \sigma(i_{B_N(t)}) - t;$$

$V_N(t) + t$ contains the times of infections that have been determined by time t but have not yet taken place. Provided that t is not too large, $B_N(t)$ is not very much smaller than $B'(t)$, and $|V'(t) \setminus V_N(t)|$ is also relatively small. This is the case if we take

$$t \;=\; t_N(u) \;:=\; \lambda^{-1}(\tfrac{1}{2}\log N + u), \tag{2.7}$$

for any fixed $u > 0$, since then $B'(t_N(u)) \sim We^u\sqrt{N}$, and hence the number of indices of $[N]$ chosen more than once in the construction of the epidemic up to this time has mean

$$N^{-1}\binom{B'(t_N(u))}{2} \;\sim\; \tfrac{1}{2}W^2e^{2u},$$

of relative order $O(N^{-1/2})$ when compared to $B'(t_N(u))$ as N becomes large; this observation is made precise later.

We now suppose that $W > 0$, and that the branching and epidemic processes have been coupled as described above up to the time $\tau_N := \tau(B', \lfloor\sqrt{N}\rfloor)$, where, for any $r > 0$, $\tau(B', r) := \inf\{t > 0\colon\, B'(t) \ge r\}$. We denote by $\mathcal{F}_{\tau_N}$ the corresponding σ-field, including the information in the sets $V'(\tau_N)$ and $V_N(\tau_N)$, but not that of the labels that are to be assigned to them for the epidemic process. Since $B'(t)e^{-\lambda t} \to W$ a.s. as $t \to \infty$, it follows that $B'(t-)/B'(t) \to 1$ a.s. also, and hence that $\lim_{N\to\infty} N^{-1/2}B'(\tau_N) = 1$ a.s. as $N \to \infty$. Thus

$$\tau_N \;=\; \lambda^{-1}\{\log B'(\tau_N) - \log W(\tau_N)\} \;\sim\; \lambda^{-1}\{\tfrac{1}{2}\log N - \log W\}$$

as $N \to \infty$. Note that $B'(\tau_N) = \lfloor\sqrt{N}\rfloor$ if G is absolutely continuous with respect to Lebesgue measure and multiple births are excluded.

We now examine whether, and if so when, a randomly chosen individual $K \in [N]$ becomes infected. To do so, we begin by writing

$$J_N \;:=\; [N] \setminus \{i_r,\, 1 \le r \le B'(\tau_N)\} \tag{2.8}$$

to denote the set of indices that have not been used in the definition of the epidemic up to time τ_N, and we set

$$J_{Nl} \;:=\; \{j \in J_N\colon\, \nu(j) = l\}, \quad M_{Nl} \;:=\; |J_{Nl}| \quad \text{and} \quad M_N \;:=\; \sum_{j\in J_N}\nu(j) \;=\; \sum_{l\ge 1} lM_{Nl}. \tag{2.9}$$

We then let

$$G_{Nl,k}(x) \;:=\; \frac{1}{M_{Nl}}\sum_{j\in J_{Nl}} I[\tau(j,k) \le x] \tag{2.10}$$

denote the empirical distribution function of the times of the k-th in order potential infections of individuals that have l such in total, and write

$$G_N(x) \;:=\; \frac{1}{M_N}\sum_{l\ge 1} M_{Nl}\sum_{k=1}^{l} G_{Nl,k}(x) \;=\; \frac{1}{M_N}\sum_{j\in J_N}\xi_j(0,x] \tag{2.11}$$

for the overall empirical distribution of the infection times of individuals in J_N. We introduce the σ-field

$$\mathcal{F}^+_{\tau_N} = \mathcal{F}_{\tau_N} \bigvee \sigma(\{\tau(j,k),\, 1 \le k \le \nu(j), j \in J_N\}). \tag{2.12}$$

If $K \in [N] \setminus J_N$, it has already been infected during the epidemic process before time τ_N; the conditional probability of this occurring is $\zeta_N := N^{-1}B'(\tau_N)$, and this is small. If not, it can only have been infected if there is a chain of infection running backwards from K to one of the $|V_N(\tau_N)|$ individuals in J_N that were infected by individuals in $[N] \setminus J_N$, but at times after τ_N. Now the M_N infection events originating from individuals in J_N are directed at independently and randomly chosen individuals in $[N]$. Hence, K is potentially directly infected as a result of a set of $\mathrm{Bi}\,(M_N, 1/N)$–many events; the individuals that infect K (its generation 1 predecessors) were themselves infected at times preceding the infection of K by amounts realized through a Bernoulli$(1/N)$ thinning of the set of M_N times $\{\tau(j,k),\, 1 \le k \le \nu(j),\, j \in J_N\}$. This procedure can be iterated to determine the predecessors in successive generations, with duplicate choices of a pair (j,k) leading to 'ghosts', as before. In this way, the susceptibility process, consisting of the chains of potential infection leading to K, can be generated from a branching process $\widehat{Z}_N$ with numbers of offspring having a binomial $\mathrm{Bi}\,(M_N, 1/N)$ distribution, and occurring at times sampled independently from G_N.

For the purposes of asymptotics, it is inconvenient to have this branching process dependent on N. With some associated error, it can be replaced with a branching process $\widehat{Z}$ that has a Poisson $\mathrm{Po}\,(\mu)$ offspring distribution, noting that

$$\mu \;:=\; \sum_{l \ge 1} l p_l \;\approx\; N^{-1} M_N \;\approx\; \frac{M_N}{|J_N|}, \tag{2.13}$$

with the birth times independently sampled from the distribution G defined in (2.1). Note that we can write

$$G = \frac{1}{\mu} \sum_{l \ge 1} p_l \sum_{k=1}^{l} G_{lk}, \tag{2.14}$$

where G_{lk} is the distribution function of the time of the k-th event in ξ_1, conditional on $\nu(1) = l$. For this branching process, we can define $\widehat{B}(t)$ to be the number of births up to time t, and conclude that, under our assumptions, by Theorem 5.4 and (3.10) of Nerman (1981),

$$\widehat{B}(t) e^{-\lambda t} \;\to\; \widehat{W} \text{ a.s.}\,, \tag{2.15}$$

for a random variable $\widehat{W}$ that satisfies $\{\widehat{W} > 0\} = \{\lim_{t\to\infty} \widehat{B}(t) = \infty\}$ a.s. Furthermore, letting

$$A(t) \;:=\; \{a_t(r)\colon\, 1 \le r \le \widehat{B}(t)\},$$

where $a_t(r) := t - \hat{\sigma}(r)$ is the age at time t of the r-th individual, it also follows that, on $\{\widehat{W} > 0\}$, by Corollary 5.6 in Nerman (1981) together with the observation from p.170 of Jagers (1975),

$$\lim_{t\to\infty} \sup_{s \ge 0} \Big| \frac{1}{\widehat{B}(t)} \sum_{r=1}^{\widehat{B}(t)} I[a_t(r) \le s] - (1 - e^{-\lambda s}) \Big| \;=\; 0 \quad \text{a.s.}\,. \tag{2.16}$$

Note that, for any $\phi \ge 0$,

$$\int_0^\infty e^{-\phi t} \mu\, G(dt) \;=\; \mathbb{E}\left(\int_0^\infty e^{-\phi t}\, \xi_i(dt) \right),$$

so that the branching processes Z and $\widehat{Z}$ indeed have the same Malthusian parameter λ. We consider this branching process run until time $t_N(u)$ as in (2.7), and we show in the next section that it represents a good enough approximation to the process of chains of potential infection to K.

Finally, we assign labels from J_N independently and at random to the individuals in the set U_N, whose birth times are the elements of $\tau_N + V_N(\tau_N)$ — these are the birth times in the forward epidemic process that have been determined by time τ_N, but have not occurred by then — and also to the set $\widehat{U}_N(u)$ composed of the distinct individuals among the $\widehat{B}(t_N(u))$ that are born before $t_N(u)$ in the reverse process. If the same label is chosen for an individual in U_N, having birth time $\tau_N + v_l$, for some $v_l \in V_N(\tau_N)$, and for an individual in $\widehat{U}_N(u)$, with birth time $\hat{\sigma}(r) \le t_N(u)$, then there is a chain of infection to K of length close to

$$\begin{aligned}\tau_N + v_l + \hat{\sigma}(r) &= \lambda^{-1}\{\log\lfloor\sqrt{N}\rfloor + \tfrac{1}{2}\log N - \log W(\tau_N) + u\} + v_l - a_{t_N(u)}(r)\\ &\sim \lambda^{-1}\{\log N - \log W + u\} + v_l - a_{t_N(u)}(r);\end{aligned}$$

the actual length is $\tau_N + v_l + \hat{\sigma}_N(r)$, where $\hat{\sigma}_N(r)$ is the birth time in the $\widehat{Z}_N$ process. If, for any such pair, $v_l \le a_{t_N(u)}(r)$, so that the length of the chain of infection is no greater than $\lambda^{-1}(\log\lfloor\sqrt{N}\rfloor + \frac{1}{2}\log N - \log W(\tau_N) + u)$, and if the r-individual is not a ghost, then K is infected before this time; that is, approximately, before time $\lambda^{-1}(\log N - \log W + u)$.

2.2 Approximating $\widehat{Z}_N$ by $\widehat{Z}$

The first step to be justified is that the branching process $\widehat{Z}_N$ with offspring numbers distributed according to the binomial $\mathrm{Bi}\,(M_N, 1/N)$ distribution and with ages at birth independently sampled from G_N, as in (2.9) and (2.11), can be replaced in our considerations by the process $\widehat{Z}$, in which the offspring numbers have the Poisson $\mathrm{Po}\,(\mu)$ distribution and the ages are sampled independently from G, as in (2.13) and (2.1). We begin by showing that the two constructions lead to the same offspring numbers, with high probability conditional on $\mathcal{F}^+_{\tau_N} \cap A_N$, at least until the first $\lfloor N^{5/8}\rfloor$ sets of progeny have been sampled; here, $A_N \in \mathcal{F}^+_{\tau_N}$ is a suitably chosen event, whose complement has small probability. The exponent $5/8$ has been chosen purely for convenience; we need an exponent exceeding $1/2$ for the argument to work.

Lemma 2.3. *Let*

$$A_N \;:=\; \{|N^{-1}M_N - (1-\zeta_N)\mu| \le N^{-7/16}\} \cap \{\zeta_N \le N^{-1/2}(\mu+1)\}; \tag{2.17}$$

then $\mathbb{P}[A_N^c] = O(N^{-1/8})$. On A_N, it is possible to construct realizations of $\widehat{Z}_N$ and $\widehat{Z}$ on the same probability space, in such a way that the numbers of offspring in the first $\lfloor N^{5/8}\rfloor$ sets of progeny in the two processes are identical with conditional probability $1 - O(N^{-1/8})$.

Proof. We begin by noting from (2.9) that, conditional on $\mathcal{F}^+_{\tau_N}$, $M_N := \sum_{j\in J_N}\nu(j)$ is a sum of $N - B'(\tau_N)$ independent and identically distributed random variables with mean μ and finite variance. Hence, by Chebyshev's inequality,

$$\mathbb{P}[|N^{-1}M_N - (1-\zeta_N)\mu| > N^{-7/16}] \;\le\; N^{-1}\mathbb{E}\{\nu(1)^2\}\,N^{7/8} \;=\; O(N^{-1/8}). \tag{2.18}$$

Then observe that

$$B'(\tau_N) \;\le\; B'(0) + \sum_{j=1}^{\lfloor\sqrt{N}\rfloor - 1} X_j, \tag{2.19}$$

where X_j denotes the number of offspring of the j-th born individual (randomly ordered in the case of simultaneous births). Hence, with $B'(0) = 1$ and since $\mu > 1$,

$$\mathbb{P}[\zeta_N \geq N^{-1/2}(\mu+1)] \;\leq\; \mathbb{P}[B'(\tau_N) \geq 1 + (\lfloor\sqrt{N}\rfloor - 1)\mu + \sqrt{N}] \;\leq\; N^{-1/2}\mathrm{Var}\,\nu(1),$$

by Chebyshev's inequality.

Now the total variation distance between $\mathrm{Bi}\,(M_N, 1/N)$ and $\mathrm{Po}\,(M_N/N)$ is at most $1/N$ (Barbour, Holst & Janson (1992), (1.23)), so that branching processes with these two offspring distributions can be coupled so as to agree until after $\lfloor N^{5/8}\rfloor$ sets of progeny have been sampled with failure probability of at most $N^{-3/8}$. Then, by considering the likelihood ratio, r independent samples from Poisson distributions with means μ and μ' can be distinguished with probability at most $d_{TV}(\mathrm{Po}\,(r\mu), \mathrm{Po}\,(r\mu')) \leq r|\mu - \mu'|/\sqrt{r\mu}$; , see for example Barbour, Holst & Janson (1992), Theorem I.1.C. Hence, if $|N^{-1}M_N - \mu| \leq N^{-7/16} + \mu\zeta_N$ and $\zeta_N < N^{-1/2}(\mu+1)$, $\lfloor N^{5/8}\rfloor$ samples from $\mathrm{Po}\,(\mu)$ and from $\mathrm{Po}\,(M_N/N)$ can be coupled so as to be identical, except on an event of probability of order $O(N^{-1/8})$. This proves the lemma. □

We now proceed to the comparison between the age distributions G_N and G. We assume henceforth that $N \geq n_1$, where

$$n_1 \;:=\; \lceil 4(1+\mu)^2\rceil, \tag{2.20}$$

so that, on A_N, $\zeta_N \leq \frac{1}{2}$, and thus $M_N \geq \frac{1}{2}N\mu$ if $N \geq n_1$. Recall the σ-field $\mathcal{F}^+_{\tau_N}$ from (2.12).

Lemma 2.4. *If $N \geq n_1$, there is an event $A^*_N \in \mathcal{F}^+_{\tau_N}$ having $\mathbb{P}[(A^*_N)^c] = O(N^{-1/8})$ such that, for suitably chosen $\varepsilon_N = O(N^{-1/6})$, we have*

$$\mathbb{P}[|G_N^{-1}(U) - G^{-1}(U)| > \psi_N \,|\, A^*_N] \;\leq\; \eta_N,$$

where $\eta_N := \psi_N + 2\varepsilon_N$ and $\psi^2_N := 2\varepsilon_N G^{-1}(1-\varepsilon_N)$, and where $U \sim \mathrm{U}[0,1]$. Note that $\psi_N + \eta_N = O(N^{-1/24})$ as $N \to \infty$ if G has finite second moment.

Proof. We begin by using the Dvoretzky–Kiefer–Wolfowitz inequality, in the form given by Massart (1990), which shows that

$$\mathbb{P}[\sqrt{M_{Nl}}\sup_x |G_{Nl,k}(x) - G_{lk}(x)| > z] \;\leq\; 2e^{-2z^2}$$

for any $z > \sqrt{\frac{1}{2}\log 2}$ and any k, l. Taking $z_N := \sqrt{2\log N}$, it follows that

$$\mathbb{P}[(A^1_{Nl,k})^c] \;\leq\; 2N^{-4} \tag{2.21}$$

for each l, k, where $A^1_{Nl,k} := \{\sqrt{M_{Nl}}\sup_x |G_{Nl,k}(x) - G_{lk}(x)| \leq z_N\} \in \mathcal{F}^+_{\tau_N}$. Observe that, for all x,

$$\begin{aligned}|G_N(x) - G(x)| \;\leq\; & \sum_{l=1}^{\lfloor N^{1/3}\rfloor}\sum_{k=1}^{l}\left\{\frac{M_{Nl}}{M_N}|G_{Nl,k}(x) - G_{lk}(x)| + |M_N^{-1}M_{Nl} - p_l|G_{lk}(x)\right\} \\ & + \frac{1}{M_N}\sum_{l>\lfloor N^{1/3}\rfloor} lM_{Nl} + \sum_{l>\lfloor N^{1/3}\rfloor} lp_l. \end{aligned}\tag{2.22}$$

Now, by the Chernoff inequalities (Theorem 2.3 in McDiarmid (1998)), we have

$$\mathbb{P}[(A^2_{Nl})^c] \;\leq\; N^{-3}, \qquad l \geq 0, \tag{2.23}$$

where $A^2_{Nl} := \{|M_{Nl} - |J_N|p_l| \le 4\log N(1 \vee \sqrt{Np_l})\} \in \mathcal{F}^+_{\tau_N}$, and, by Markov's inequality,

$$\mathbb{P}\big[(A^3_N)^c\big] \;\le\; N^{1/6} \sum_{l>\lfloor N^{1/3}\rfloor} lp_l \;\le\; N^{-1/6}\mathbb{E}\{\nu(1)^2\}, \tag{2.24}$$

where $A^3_N := \{\sum_{l>\lfloor N^{1/3}\rfloor} lM_{Nl} \le N^{5/6}\} \in \mathcal{F}^+_{\tau_N}$.

Define

$$A^*_N \;:=\; A_N \cap \left\{ \bigcap_{l=1}^{\lfloor N^{1/3}\rfloor} \bigcap_{k=1}^{l} A^1_{Nl,k} \right\} \cap \left\{ \bigcap_{l=1}^{\lfloor N^{1/3}\rfloor} A^2_{Nl} \right\} \cap A^3_N\,;$$

then $\mathbb{P}[(A^*_N)^c] = O(N^{-1/8})$, by Lemma 2.3, (2.21), (2.23) and (2.24). On A^*_N, from (2.22), for all $x \ge 0$, we have

$$\begin{aligned} |G_N(x) - G(x)| \;&\le\; \sum_{l=1}^{\lfloor N^{1/3}\rfloor} l\left\{ \frac{\sqrt{M_{Nl}}}{M_N}\sqrt{2\log N} + 4M_N^{-1}\log N\{1 \vee \sqrt{Np_l}\}\right\} \\ &\quad + \left|1 - \mu\frac{|J_N|}{M_N}\right| + \frac{N^{5/6}}{M_N} + N^{-1/3}\mathbb{E}\{\nu(1)^2\} \\ &=:\; \varepsilon_N \;=\; O(N^{-1/6}). \end{aligned}$$

To justify the order of the bound, note first that, from (2.23), on A^*_N,

$$M_{Nl} \;\le\; \begin{cases} 8\log N, & Np_l < 1; \\ 8\log N\sqrt{Np_l}, & 1 \le Np_l < \{4\log N\}^2; \\ 2Np_l, & Np_l > \{4\log N\}^2, \end{cases}$$

and then $\sum_{l\le\lfloor N^{1/3}\rfloor} l \le N^{2/3}$, $M_N \ge \frac{1}{2}N\mu$ for $N \ge n_1$ on A_N and, by Cauchy–Schwarz,

$$\sum_{l\le\lfloor N^{1/3}\rfloor} l\sqrt{p_l} \;\le\; \sqrt{N^{1/3}\mathbb{E}\{\nu(1)^2\}}.$$

Finally,

$$\left|1 - \mu\frac{|J_N|}{M_N}\right| \;=\; \frac{N}{M_N}\,|N^{-1}M_N - \mu(1-\zeta_N)| \;\le\; 2\mu^{-1}N^{-7/16}$$

on A_N, for $N \ge n_1$.

Now, since $G(x) - \varepsilon_N \le G_N(x) \le G(x) + \varepsilon_N$ for all $x \ge 0$, it also follows for all y that $G^{-1}(y - \varepsilon_N) \le G_N^{-1}(y) \le G^{-1}(y + \varepsilon_N)$, and thus that

$$|G_N^{-1}(y) - G^{-1}(y)| \le G^{-1}(y + \varepsilon_N) - G^{-1}(y - \varepsilon_N). \tag{2.25}$$

Hence it follows that, for any $\eta > 0$,

$$\begin{aligned} \int_0^{1-\eta} |G_N^{-1}(y) - G^{-1}(y)|\,dy \;&\le\; \int_0^{1-\eta} \{G^{-1}(y + \varepsilon_N) - G^{-1}(y - \varepsilon_N)\}\,dy \\ &\le\; \int_{1-\eta-\varepsilon_N}^{1-\eta+\varepsilon_N} G^{-1}(y)\,dy \;\le\; 2\varepsilon_N G^{-1}(1 - \eta + \varepsilon_N) \;=\; \psi_N^2. \end{aligned}$$

Taking $\eta := 2\varepsilon_N$, this shows that, for U uniformly distributed on $[0,1]$,

$$\mathbb{E}\{|G_N^{-1}(U) - G^{-1}(U)|I[U \le 1 - 2\varepsilon_N]|A^*_N\} \;\le\; \psi_N^2,$$

and Markov's inequality completes the proof. Note that, since G is assumed to have finite second moment, $x^2(1 - G(x)) = o(1)$ as $x \to \infty$, implying that $\varepsilon_N G^{-1}(1 - \varepsilon_N) = o(\varepsilon_N^{1/2})$ as $N \to \infty$. □

Corollary 2.5. *Let A_N^* be as in Lemma 2.4. If G satisfies Assumption 2 with $b < \infty$, then, on A_N^*,*

$$\sup_{0\le u\le 1} |G_N^{-1}(u) - G^{-1}(u)| \;=\; o(1) \;\text{ as } N\to\infty;$$

if G satisfies Assumption 2 with $b = \infty$, then

$$\sup_{u:\; G^{-1}(u)\le x_N} |G_N^{-1}(u) - G^{-1}(u)| \;=\; o(1) \;\text{ as } N\to\infty,$$

for x_N such that $N^{-\alpha}x_N$ is bounded below as $N\to\infty$ for some $\alpha > 0$.

Proof. For the first part, let the support of G be $[a, b]$. Then for any $\delta > 0$, with (2.25),

$$|G_N^{-1}(u) - G^{-1}(u)| \;\le\; 2\delta + \frac{2\varepsilon_N}{g_-[a+\delta, b-\delta]},$$

where $\varepsilon_N = O(N^{-1/6})$ is as in Lemma 2.4, and $g_-[c, d] := \inf_{c\le x\le d} g(x)$. So take $\delta = \delta_N \to 0$ in such a way that $\varepsilon_N = o(g_-[a+\delta_N, b-\delta_N])$.

For the second part, for N large enough that $k(x_0+1)^{-\gamma} > \varepsilon_N$, define $x_{N1} > x_0$ such that $(x_{N1}+1)^\gamma = k/\varepsilon_N$, and choose any $x_N \le x_{N1}$. Then, uniformly for all u such that $a+\delta \le G^{-1}(u) \le x_N$,

$$G^{-1}(u+\varepsilon_N) - G^{-1}(u) \;\le\; \frac{\varepsilon_N}{\min\{g_-[a+\delta, x_0], k(x_N+1)^{-\gamma}\}} \;\le\; \frac{\varepsilon_N}{g_-[a+\delta, x_0]} + \frac{\varepsilon_N}{k(x_N+1)^{-\gamma}}.$$

So choose $\delta_N = \varepsilon_N^{1/2}$ and $x_N = (k\delta_N/\varepsilon_N)^{1/\gamma} - 1 \le x_{N1}$, and δ'_N such that $\varepsilon_N = o(g_-[a+\delta'_N, x_0])$; this gives

$$\sup_{u:G^{-1}(u)\le x_N} |G_N^{-1}(u) - G^{-1}(u)| \;\le\; \frac{2\varepsilon_N}{g_-[a+\delta'_N, x_0]} + 2\delta_N + \delta'_N \;\to\; 0,$$

and $x_N N^{-1/(12\gamma)}$ is bounded below as $N\to\infty$. □

We also need to know that paths of a given length cannot contain too many births.

Lemma 2.6. *Suppose that $\lim_{\varepsilon\to 0} G(\varepsilon) = 0$. Then there exist $t_* > 0$ such that all individuals of generation n in $\widehat{Z}$ are born after time nt_*, except on an event of probability at most $2e^{-n}$.*

Proof. Let $\widehat{Z}_n$ denote the number of individuals of generation n in $\widehat{Z}$, starting with $\widehat{Z}_0 = 1$. Then $\mathbb{E}\widehat{Z}_n = \mu^n$, and so $\mathbb{P}[\widehat{Z}_n > \{e\mu\}^n] \le e^{-n}$. Now the time elapsed up to generation n along any given line is a sum of n independent G–distributed random variables, and the probability that fewer than $n/2$ of these are greater than a given value ε is the binomial probability

$$\mathrm{Bi}\,(n,p)[\lceil n/2\rceil, n] \;\le\; \{1+p(z_p-1)\}^n z_p^{-\lceil n/2\rceil} \;\le\; (4p)^{n/2}\,,$$

with $z_p := (1-p)/p$ and $p = G(\varepsilon)$. Hence the probability that, up to generation n, any line takes less that time $\varepsilon n/2$ is at most

$$e^{-n} + \exp\{n(\log\mu + 1) - \tfrac{1}{2}n\log(1/4G(\varepsilon))\}.$$

Taking $\varepsilon > 0$ such that $\log(1/4G(\varepsilon)) \ge 2(\log\mu + 2)$ makes this probability at most $2e^{-n}$, and taking $t_* := \varepsilon/2$ proves the lemma. □

2.3 Controlling the ghosts

We now need to control the differences between the epidemic and branching processes; we need to show that ghosts play no significant part. We begin with the forward branching process Z. Recalling from (2.4) that $W(t) := B'(t)e^{-\lambda t} \to W$ a.s., we write $e_W := \sup_t \mathbb{E}W(t) < \infty$. Label the individuals of Z independently and uniformly from $[N]$ in order of birth epoch until time τ_N; let $L(t)$ denote the number of times that a label has been used before, creating an initial ghost, and let $L_+(t) \geq L(t)$ denote the number of initial ghosts and their descendants whose birth times have been determined by time t. Finally, let $t_N^\alpha := \alpha\lambda^{-1}\log N$, $\alpha > 0$.

Lemma 2.7. *Under the above assumptions,*

$$\mathbb{P}[\{N^{-1/2}L_+(\tau_N) \geq N^{-1/4}\} \cap \{W(\tau_N) \geq N^{-1/8}\}] \;=\; O(N^{-1/8}\log N).$$

Proof. For any of the first $\lfloor\sqrt{N}(\mu+1)\rfloor$ indices chosen, the probability that it is a repeat of an index chosen earlier is at most $N^{-1/2}(\mu+1)$. Hence, for any $\alpha > 0$, writing $T := t_N^\alpha$,

$$\mathbb{E}\{L_+(\tau_N \wedge t_N^\alpha)\} \quad\leq\quad (\mu+1)N^{-1/2}\mathbb{E}\left\{\int_0^T \mu e_W e^{\lambda(T-t)}\, B(dt)\right\},$$

since an individual born at t has an expected number of descendants at time T of at most $e_W e^{\lambda(T-t)}$, for each of which the expected number of offspring whose births are still to come is at most μ. Hence

$$\begin{aligned}\mathbb{E}\{L_+(\tau_N \wedge t_N^\alpha)\} &\leq (\mu+1)N^{-1/2}\mu e_W e^{\lambda T}\mathbb{E}\left\{B(T)e^{-\lambda T} + \lambda\int_0^T e^{-\lambda t}B(t)\,dt\right\}\\ &\leq (\mu+1)N^{-1/2}\mu e_W^2(1+\lambda T)e^{\lambda T}.\end{aligned}$$

Thus, choosing $\alpha = (1+\varepsilon)/2$, we have

$$\mathbb{P}[\{N^{-1/2}L_+(\tau_N) \geq N^{-1/2+\varepsilon}\} \cap \{W(\tau_N) \geq N^{-\varepsilon/2}\}] \;=\; O(N^{-\varepsilon/2}\log N),$$

since $\tau_N \leq t_N^{(1+\varepsilon)/2}$ when $W(\tau_N) \geq N^{-\varepsilon/2}$, and the lemma follows by taking $\varepsilon = 1/4$. □

For the backward branching processes $\widehat{Z}_N$ and $\widehat{Z}$, the argument is a little different, because the identities of the individuals (even if not their labels) are implicitly recognised during the construction of the branching process $\widehat{Z}_N$; the choice of a particular value from G_N may well determine the choice of the individual in J_N that gave rise to it, and will certainly do so if the distribution G is continuous. Hence, when constructing $\widehat{Z}_N$, an initial ghost appears when the same birth time $t_{Nj,l}$ is sampled from the same individual j for the second or subsequent time, and individual j is represented more than once (but without creating ghosts) if several distinct elements of $\{t_{Nj,l},\, 1 \leq l \leq j\}$ are sampled. By Lemma 2.3, the branching process $\widehat{Z}$ has the same offspring numbers as $\widehat{Z}_N$ up to $\lfloor N^{5/8}\rfloor$ with probability $1 - O(N^{-1/8})$, and individuals can also be identified starting from a realization of $\widehat{Z}$, by using the quantile transformation to go from a value sampled from G to the corresponding value from G_N (with an arbitrary rule for distinguishing individuals that give rise to identical birth times). Thus the ghosts arise during the joint construction; afterwards, labelling is at random without replacement from J_N for the *distinct* individuals in $\widehat{Z}$ up to time $\lfloor N^{5/8}\rfloor$.

As before, we note that $\widehat{W}(t) := \widehat{B}(t)e^{-\lambda t} \to \widehat{W}$ a.s. as $t \to \infty$. We can then write $e_{\widehat{W}} := \sup_t \mathbb{E}\{\widehat{W}(t) \,|\, \widehat{B}(0) = 1\} < \infty$ (if the process is started with $\widehat{B}(0) = 2$, as from K and K', the supremum is doubled). We let $\widehat{L}(t)$ denote the number of initial ghosts that

have arisen by time t, $\widehat{L}_+(t) \geq \widehat{L}(t)$ the number of initial ghosts and their descendants that have arisen by then, and $\widetilde{L}^{(2)}(t)$ the number of individuals represented at least twice by time t. We also denote by $\widehat{L}^\theta(t)$ the number of marked individuals and their descendants up to time t, if individuals are marked independently with probability θ.

Lemma 2.8. *Let K and K' be independently chosen at random from J_N, and let $\eta'_N := \eta_N \log N$, where $\eta_N = o(N^{-1/24})$ is as in Lemma 2.4. Then, conditional on A^*_N, and starting the branching process $\widehat{Z}$ either from K or from both of K and K', we have*

$$(1) \qquad \mathbb{P}[N^{-1/2}\widehat{L}_+(t_N(u)) \geq N^{-3/16} \,|\, \mathcal{F}^+_{\tau_N} \cap A_N] \;=\; O(N^{-1/8}\log N);$$
$$(2) \qquad \mathbb{P}[N^{-1/2}\widehat{L}^{\theta(N)}(t_N(u)) \geq N^{-3/16}] \;=\; O(\theta(N)N^{5/24}\log N),$$

uniformly for all $u \leq (\log N)/48$. Furthermore, there is a set $A^4_N \in \mathcal{F}^+_{\tau_N}$ with $\mathbb{P}[(A^4_N)^c] = O(N^{-1/24})$ such that

$$(3) \qquad \mathbb{P}[N^{-1/2}\widetilde{L}^{(2)}(t_N(u)) \geq N^{-7/24} \,|\, \mathcal{F}^+_{\tau_N} \cap A^4_N] \;=\; O(N^{-1/24}),$$

uniformly in the same range of u.

Proof. The first and second statements of the lemma are proved in much the same way as Lemma 2.7. For the first, we note that the probability of the r-th individual born being an initial ghost is at most $(r-1)/M_N$. Hence, for any $w > 0$ and $N \geq n_1$,

$$\begin{aligned}\mathbb{E}\big\{\min\{\widehat{L}_+(t_N(u), \bar{\tau}(\widehat{B}, we^{\lambda t_N(u)}))\} \,|\, \mathcal{F}^+_{\tau_N} \cap A_N\big\} \\ &\leq\; M_N^{-1}N^{1/2}we^u\mathbb{E}\left\{\int_0^{t_N(u)} e_W e^{\lambda(t_N(u)-t)}\,\widehat{B}(dt)\right\}, \\ &\leq\; 2\mu^{-1}we^{2u}e_W^2(1+u+\tfrac{1}{2}\log N),\end{aligned}$$

where $\bar{\tau}(\widehat{B}, v) := \inf\{t\colon\ \widehat{B}(t) \geq v\}$. Thus, and from Lemma 2.3,

$$\begin{aligned}\mathbb{P}[\{N^{-1/2}\widehat{L}_+(t_N(u)) \geq N^{-1/2+5\varepsilon/4}\} \cap \{\bar{\tau}(\widehat{B}, N^{\varepsilon/2}e^{\lambda t_N(u)}) > t_N(u)\} \,|\, \mathcal{F}^+_{\tau_N} \cap A_N] \\ &=\; O(N^{-3\varepsilon/4}e^{2u}\log N).\end{aligned}$$

Since also

$$\mathbb{P}[\bar{\tau}(\widehat{B}, N^{\varepsilon/2}e^{\lambda t_N(u)}) \leq t_N(u)] \;=\; \mathbb{P}[\widehat{B}(t_N(u)) \geq N^{\varepsilon/2}e^{\lambda t_N(u)}] \leq N^{-\varepsilon/2}e_{\widehat{W}},$$

it follows that, for $u \leq \frac{1}{8}\varepsilon \log N$,

$$\mathbb{P}[N^{-1/2}\widehat{L}_+(t_N(u)) \geq N^{-1/2+5\varepsilon/4} \,|\, \mathcal{F}^+_{\tau_N} \cap A_N] \;=\; O(N^{-\varepsilon/2}\log N),$$

and the first statement follows by taking $\varepsilon = 1/4$.

For the second, we have

$$\begin{aligned}\mathbb{E}\{N^{-1/2}\widehat{L}^\theta(t_N(u))\} \;&\leq\; N^{-1/2}\theta\mathbb{E}\left\{\int_0^{t_N(u)} e_{\widehat{W}}e^{\lambda(t_N(u)-t)}\,\widehat{B}(dt)\right\}, \\ &\leq\; \theta e^u e^2_{\widehat{W}}(1+u+\tfrac{1}{2}\log N),\end{aligned}$$

and the statement follows from Markov's inequality.

For the third, we begin by noting that the choices of individual in $\widehat{Z}_N$ after n have been examined are multinomially $\mathrm{MN}\big(n; \{\nu(j)/M_N,\, j \in J_N\}\big)$ distributed, so that the mean number of individuals that have by then been chosen more than once is at most

$$\frac{n^2}{2}\sum_{j\in J_N}\left(\frac{\nu(j)}{M_N}\right)^2 \;\leq\; \frac{n^2}{2}\sum_{l\geq 1} M_{Nl}\left(\frac{l}{M_N}\right)^2. \tag{2.26}$$

Let $A_N^4 := \{\sum_{l\geq 1} M_{Nl}(l/M_N)^2 \leq 2N^{\varepsilon-1}\} \in \mathcal{F}^+_{\tau_N}$, and suppose that $N \geq n_1$ as in (2.20). Observe that, since $M_N \geq \frac{1}{2}N\mu$ on A_N, and

$$\mathbb{E}\left\{\sum_{l\geq 1} M_{Nl}\left(\frac{2l}{N\mu}\right)^2\right\} \;\leq\; \sum_{l\geq 1} Np_l\left(\frac{2l}{N\mu}\right)^2 \;\leq\; 4N^{-1}\mathbb{E}\{\nu(1)^2\}\mu^{-2},$$

we have $\mathbb{P}[(A_N^4)^c] = O(N^{-\varepsilon})$ for any $\varepsilon < 1/8$. Then, using (2.26),

$$\mathbb{E}\{\widetilde{L}^{(2)}(t)I[\widehat{W}(t) \leq N^\varepsilon] \,|\, A_N^4\} \;=\; \mathbb{E}\{\widetilde{L}^{(2)}(t)I[\widehat{B}(t) \leq N^\varepsilon e^{\lambda t}] \,|\, A_N^4\} \;\leq\; N^{-1+3\varepsilon}\, e^{2\lambda t} \;\leq\; N^{4\varepsilon},$$

uniformly in $t \leq (1/2\lambda)(1+\varepsilon)\log N$. Hence, and since $\mathbb{P}[\widehat{W}(t) > N^\varepsilon] \leq e_{\widehat{W}}N^{-\varepsilon}$, it follows that, for $u \leq \frac{1}{2}\varepsilon \log N$,

$$\mathbb{P}[N^{-1/2}\widetilde{L}^{(2)}(t_N(u)) \geq N^{5\varepsilon-1/2} \,|\, A_N^4] \;=\; O(N^{-\varepsilon}),$$

giving the third assertion if we take $\varepsilon = 1/24$. □

We now use $\widehat{L}(G_N, t)$ to denote the number of individuals in $\widehat{Z}$, together with their descendants, up to time t, for which the sample taken from G to determine their birth time is such that the difference between it and the corresponding value obtained from G_N by the quantile transformation exceeds the threshold ψ_N defined in Lemma 2.4. Note that, on A_N^*, the expected contribution to $\widehat{L}(G_N, t)$ resulting from the offspring of an individual born at time $v < t$ is at most $\mu\eta_N e_{\widehat{W}} e^{\lambda(t-v)}$, where η_N is as in Lemma 2.4. The proof of Lemma 2.8(2) then yields the following corollary.

Corollary 2.9. *In the setting of Lemma 2.8, with $\eta'_N = \eta_N \log N$ and η_N as in Lemma 2.4, we have*

$$\mathbb{P}[N^{-1/2}\widehat{L}(G_N, t_N(u)) \geq (\eta'_N)^{1/2} \,|\, \mathcal{F}^+_{\tau_N} \cap A_N^*] \;=\; O((\eta'_N)^{1/2}).$$

2.4 Main theorem

We now combine our previous results to prove the main result of Section 2. For any $t \geq 0$, let $\mathcal{S}_N(t)$ denote the set of individuals in the epidemic that are still susceptible at time t, and write $S_N(t) := |\mathcal{S}_N(t)|$. Then, for independently and randomly chosen K and K' in $[N]$,

$$\mathbb{E}\{N^{-1}S_N(t) \,|\, \mathcal{F}^+_{\tau_N}\} \;=\; \frac{1}{N}\sum_{k=1}^N \mathbb{P}[k \in \mathcal{S}_N(t) \,|\, \mathcal{F}^+_{\tau_N}] \;=\; \mathbb{P}[K \in \mathcal{S}_N(t) \,|\, \mathcal{F}^+_{\tau_N}],$$

and similarly

$$\mathrm{Var}\,\{N^{-1}S_N(t) \,|\, \mathcal{F}^+_{\tau_N}\} \;=\; \mathbb{P}[\{K, K'\} \subset \mathcal{S}_N(t) \,|\, \mathcal{F}^+_{\tau_N}] - \{\mathbb{P}[K \in \mathcal{S}_N(t) \,|\, \mathcal{F}^+_{\tau_N}]\}^2,$$

and we use these expressions to show that $N^{-1}S_N(t)$ is close to its expectation, and to give an asymptotic expression for it.

At time τ_N, the epidemic process has generated a collection U_N of individuals, whose birth times, the elements of $V_N(\tau_N)$, are determined, but have not yet occurred, and which have not yet been labelled (so that some of them may turn out to be ghosts); labels are assigned to them independently and at random from $[N]$, and ghosts are then removed, leaving a labelled set $U'_N \subset U_N$.

A randomly chosen individual K in $[N]$ samples an independent copy of the reversed branching process $\widehat{Z}$, and uses it to determine its susceptibility process, by way of $\widehat{Z}_N$.

For times to infection, as measured in $\widehat{Z}$-time, not exceeding $t_N(u+h)$, there is a corresponding susceptibility set $\widehat{U}_N(u+h)$, consisting of distinct individuals. Labels are now assigned to the elements of the set $\widehat{U}_N(u+h)$, chosen independently but *without replacement* from J_N. Let $\mathcal{E}_N(u+h)$ denote the set of elements of $\widehat{U}_N(u+h)$ that share labels with members of U_N. Then $E_N(u+h) := |\mathcal{E}_N(u+h)|$ has conditional expectation $|\widehat{U}_N(u+h)|\,|U_N|/N$. If $E_N(u+h) = 0$, there is no path of infection from i_1 to K of $\widehat{Z}$-length less than $\tau_N + u + h$. If $E_N(u+h) > 0$, go through the elements of $\mathcal{E}_N(u+h)$ in order of increasing $\widehat{Z}$-time, and mark all their progeny in $\widehat{U}_N(u+h)$ as ghosts, since these elements are also represented as members of U_N, and their infection pre-history has already been determined in $\mathcal{F}^+_{\tau_N}$. Let $\mathcal{E}'_N(u+h) \subset \mathcal{E}_N(u+h)$ denote those elements of $\mathcal{E}_N(u+h)$ that are not marked as ghosts, and write $E'_N(u+h) := |\mathcal{E}'_N(u+h)|$. For any element e of $\mathcal{E}'_N(u+h)$, let $\tau_N + v$ denote the birth time of the corresponding element of U'_N, let $\hat{\sigma}$ denote the birth time in $\widehat{Z}$ of the element of $\mathcal{E}'_N(u+h)$, and $\hat{\sigma}_N$ its corresponding birth time in $\widehat{Z}_N$. Then e gives rise to an infection path from i_1 to K of length $\tau_N + v + \hat{\sigma}_N$. If this is less than or equal to $\tau_N + t_N(u)$ for any e, then $K \notin \mathcal{S}_N(\tau_N + t_N(u))$; otherwise, $K \in \mathcal{S}_N(\tau_N + t_N(u))$ unless, possibly, there is an infection path with $v + \hat{\sigma}_N \le t_N(u)$ but $v + \hat{\sigma} > t_N(u+h)$. Using these considerations, we can deduce an approximation for $\mathbb{P}[K \in \mathcal{S}_N(\tau_N + t_N(u)) \,|\, \mathcal{F}^+_{\tau_N}]$, and a similar argument, with two reversed branching processes, leads also to a corresponding approximation to $\mathbb{P}[\{K, K'\} \subset \mathcal{S}_N(\tau_N + t_N(u))]$.

The proof of the theorem that follows is essentially concerned with quantifying the above steps. In particular, it is to be shown that $\mathcal{E}_N(u+h) = \mathcal{E}'_N(u+h)$ with high probability, and that $|\widehat{U}_N(u+h)|\,|U_N|/N \sim (\mu-1)\widehat{W}e^{u+h}$. Then, for any element e of $\mathcal{E}'_N(u+h)$, we need to show that the corresponding v is sampled from a distribution close to F, as defined in (2.6), and that $t_N(u+h) - \hat{\sigma}_N$ is sampled from a distribution close to the exponential distribution Exp(λ) with mean $1/\lambda$, in view of (2.16). Assuming that this is the case, it follows that

$$\mathbb{P}[v + \hat{\sigma}_N \le t_N(u)] \;\sim\; e^{-h}\int_0^\infty \lambda e^{-\lambda s} F(s)\,ds \;=\; e^{-h}\frac{\mu}{\mu-1}\int_0^\infty \lambda s e^{-\lambda s}\,G(ds). \quad (2.27)$$

The conditional mean number of such events is therefore asymptotically $\widehat{W}e^u m_*$, where m_* is given in (2.3). and a Poisson approximation shows that the probability of none of them occurring is close to $e^{-\widehat{W}e^u m_*}$. The required approximation to $\mathbb{P}[K \in \mathcal{S}_N(\tau_N + t_N(u)) \,|\, \mathcal{F}^+_{\tau_N}]$ is then $\mathbb{E}\{e^{-\widehat{W}e^u m_*}\}$. Finally, the possibility that there is an infection path with $v + \hat{\sigma}_N \le t_N(u)$ but $v + \hat{\sigma} > t_N(u+h)$ has to be excluded.

Theorem 2.10. *Under Assumptions 1 and 2, there exists an event $\widetilde{A}_N \in \mathcal{F}^+_{\tau_N}$ such that $\mathbb{P}[\widetilde{A}^c_N] \to 0$ as $N \to \infty$, for which*

$$\mathbb{P}\Big[\sup_u |N^{-1}S_N(\tau_N + \lambda^{-1}\{\tfrac{1}{2}\log N + u\}) - \hat{s}(u)| > \varepsilon \,\Big|\, \mathcal{F}^+_{\tau_N} \cap \widetilde{A}_N \cap \{\tau_N < \infty\}\Big] \;\to\; 0$$

as $N \to \infty$, where $\hat{s}$ is the decreasing function given by

$$\hat{s}(u) \;:=\; \mathbb{E}\big\{\exp\{-\widehat{W}e^u m_*\}\big\},$$

and where $m_* = \mu\lambda\int_0^\infty s e^{-\lambda s}\,G(ds)$, as in (2.3).

Remark 2.11. *It therefore follows that $\sup_u |N^{-1}S_N(\lambda^{-1}\{\log N - \log W + u\}) - \hat{s}(u)| \to_d 0$, conditionally on $W > 0$. However, in practice, it may be more reasonable to expect to be able to observe the time τ_N than it is to know the value of W, or, equivalently, when the first infection occurred.*

Proof. By Lemma 2.7, and by Corollary 5.6 in Nerman (1981) with $\phi_1(t) = \xi(t,\infty)$ and $\phi_2(t) = 1$, and using the fact that $N^{-1/2}B'(\tau_N) \to 1$ a.s. as $N \to \infty$, we obtain that

$N^{-1/2}|U_N| \to (\mu-1)$ a.s. as $N \to \infty$ on $\{W > 0\}$; Lemma 2.7 shows that excluding ghosts has negligible effect on the branching asymptotics. Thus we can define a set

$$A_N^5 \;:=\; \{|N^{-1/2}|U_N| - (\mu-1)\,| \le \eta_1(N)\} \;\in\; \mathcal{F}^+_{\tau_N}, \tag{2.28}$$

where $\eta_1(N) \to 0$ and $\mathbb{P}[(A_N^5)^c] \to 0$ as $N \to \infty$. Let

$$\widetilde{A}_N \;:=\; A_N^* \cap A_N^4 \cap A_N^5 \cap \{W(\tau_N) \ge N^{-1/8}\}.$$

We wish first to show that, for any $u \in \mathbb{R}$,

$$\mathbb{P}[K \in \mathcal{S}(\tau_N + t_N(u))\,|\,\mathcal{F}^+_{\tau_N} \cap \widetilde{A}_N] \;\sim\; \hat{s}(u),$$

where $\hat{s}$ is as stated in the theorem. To do so we proceed as outlined above. On $\widetilde{A}_N$, we have $|U_N| \sim N^{1/2}(\mu-1)$, in view of (2.28). Then, by (2.15) and Lemma 2.8(1,3), $|\widehat{U}_N(u+h)| \sim N^{1/2}e^{(u+h)}\widehat{W}$; Lemma 2.8 shows that excluding ghosts and individuals multiply referenced has little effect on the branching asymptotics. The mean number of individuals in $\widehat{U}_N(u+h)$ that share a common index with a member of U_N is thus asymptotic to

$$N^{1/2}(\mu-1).N^{1/2}\widehat{W}e^{(u+h)}/N \;=\; \widehat{W}(\mu-1)e^{(u+h)}.$$

We now show that $\mathbb{P}[\mathcal{E}_N(u+h) \ne \mathcal{E}'_N(u+h)] = O(N^{-3/16})$. Letting $E_N^D(u+h)$ denote the number of descendants of $\mathcal{E}_N(u+h)$, it follows from Lemma 2.8(2), by taking $\theta = \theta(N) = N^{-1}|U_N|$ and in view of (2.28), that

$$\mathbb{P}[E_N^D(u+h) \ge N^{5/16}\,|\,\mathcal{F}^+_{\tau_N} \cap \widetilde{A}_N] \;=\; O(N^{-1/4}\log N).$$

The conditional probability that any of them is marked by a label from U_N is thus at most of order $O(N^{-1/2+5/16} + N^{-1/4}\log N) = O(N^{-3/16})$.

Now, because of the random scheme of assignment of labels, any pair in $\mathcal{E}_N(u+h)$ is associated with a random choice of elements v of $V_N(\tau_N)$ and a of $A(t_N(u+h))$, and the empirical distributions of the elements of these sets converge, as observed in (2.5), (2.6) and (2.16). Furthermore, the empirical distribution $\widehat{F}_N^{(u+h)}$ of the birth times in $\widehat{Z}_N$ corresponding to the elements of $A(t_N(u+h))$ also converges to the exponential Exp(λ) distribution with mean $1/\lambda$ if $\widehat{W} > 0$. To see this, we argue as follows. Recalling (2.16), let $\eta_2(t)$ be such that $\lim_{t\to\infty}\eta_2(t) = 0$ and that

$$\mathbb{P}\left[\sup_{s\ge 0}\Big|\frac{1}{\widehat{B}(t)}\sum_{r=1}^{\widehat{B}(t)} I[a_t(r) \le s] - (1-e^{-\lambda s})\Big| > \eta_2(t)\;\Big|\;\widehat{W} > 0\right] \;\le\; \eta_2(t). \tag{2.29}$$

Then define

$$k \;:=\; \left\lceil \frac{1+\varepsilon}{2\lambda t_*} \right\rceil,$$

where t_* is as in Lemma 2.6. Observe that, in view of Corollary 2.9 and of Lemma 2.6,

$$\sup_s |\widehat{F}_N^{(u)}(s) - (1-e^{-\lambda s})| \;\le\; \lambda\psi_N k\log N + \eta_2(t_N(u)) + N^{1/2}(\eta'_N)^{1/2}/|A(t_N(u))|,$$

on $\{\widehat{W} > 0\}$, uniformly in $u \le \frac{1}{2}\varepsilon\log N$, except on a set of conditional probability at most $(\eta'_N)^{1/2} + 2N^{-(1+\varepsilon)/2\lambda t_*} + \eta_2(t_N(u))$, and that $\sup_t e^{-\lambda t}|A(t)| < \infty$.

At this point, we also need to exclude the possibility that there is an infection path with $v+\hat{\sigma}_N \le t_N(u)$ but $v+\hat{\sigma} > t_N(u+h)$. Corollary 2.9 shows that, on A_N^*, the probability of having a path from K to U_N containing a sample $\tilde{\tau}$ from G such that $|\tilde{\tau} - \tilde{\tau}_N| > \psi_N$, where $\tilde{\tau}_N := G_N^{-1}(G(\tilde{\tau}))$, before time $\lambda^{-1}(1/2+\varepsilon)\log N$ is small for $\varepsilon < 1/24$, and the

number of births in a path up to that time is bounded by $c \log N$ in view of Lemma 2.6, with high probability. Hence there has to be at least one pair $(\tilde{\tau}, \tilde{\tau}_N)$ in the path such that $\tilde{\tau} - \tilde{\tau}_N > c'$, for $c' = (1/2c\lambda)\varepsilon$, if $u < \frac{1}{2}\lambda^{-1}\varepsilon \log N$ and $\hat{\sigma} - \hat{\sigma}_N \geq \lambda^{-1}\varepsilon \log N - u$. But this cannot be the case, for N large enough, in view of Corollary 2.5.

Hence, on the event $\widetilde{A}_N$, and conditional on $\mathcal{F}_N(u) := \sigma(\widehat{Z}(t), 0 \leq t \leq t_N(u)) \bigvee \mathcal{F}^+_{\tau_N}$, the mean number of pairs with common index, one from U_N and one from $\widehat{U}_N(u+h)$, that are not ghosts and give rise to an infection path between i_1 and K of length at most $\tau_N + t_N(u)$, is given as in (2.27) and (2.3) by

$$m_N(u, \widehat{W}) \quad \sim \quad \widehat{W} e^u m_*;$$

of course, the asymptotics are valid also when $\widehat{W} = 0$. Let $I_{Nj}(u)$ denote the indicator of the event that the label of the j-th element of U_N is matched with one of the labels assigned to $\widehat{U}_N(u)$, $1 \leq j \leq |U_N|$. Then, conditional on $\mathcal{F}^{(u)}_N$, $(I_{Nj}(u), 1 \leq j \leq |U_N|)$ is a collection of independent indicator random variables, each with probability $p_N(u) := |U_N|^{-1} m_N(u, \widehat{W})$; hence it follows by Barbour, Holst & Janson (1992, (1.23)) that

$$\left| \mathbb{P}\Big[\sum_{j=1}^{|U_N|} I_{Nj}(u) = 0 \,\Big|\, \mathcal{F}^{(u)}_N \cap \widetilde{A}_N \Big] - \exp\{-m_N(u, \widehat{W})\} \right| \;\leq\; p_N(u). \tag{2.30}$$

Thus we deduce that

$$\begin{aligned} &\mathbb{P}[K \in \mathcal{S}_N(\tau_N + t_N(u)) \,|\, \mathcal{F}^+_{\tau_N} \cap \widetilde{A}_N] \\ &\quad \sim\; \mathbb{E}\big\{\exp\{-m_N(u, \widehat{W})\} \,|\, \mathcal{F}^+_{\tau_N} \cap \widetilde{A}_N\big\} \;\sim\; \mathbb{E}\big\{\exp\{-\widehat{W} e^u m_*\}\big\} \;=\; \hat{s}(u). \end{aligned}$$

But this means that

$$\begin{aligned} \hat{s}_N(u) \quad &:= \quad \mathbb{E}\big\{N^{-1} S_N(\tau_N + \lambda^{-1}\{\tfrac{1}{2}\log N + u\}) \,|\, \mathcal{F}^+_{\tau_N} \cap \widetilde{A}_N\big\} \\ &= \quad \mathbb{P}[K \in \mathcal{S}_N(\tau_N + t_N(u)) \,|\, \mathcal{F}^+_{\tau_N} \cap \widetilde{A}_N] \;\sim\; \hat{s}(u) \end{aligned} \tag{2.31}$$

also.

The argument for approximating the probability that both points K and K' belong to $\mathcal{S}_N(\tau_N + t_N(u))$ runs in much the same way. The limiting random variable for the branching process $\widehat{Z}$ starting with two individuals can be expressed as $\widehat{W}_1 + \widehat{W}_2$, where the two are independent copies of $\widehat{W}$, and the sizes of the corresponding sets $\widehat{U}^{(1)}_N(u)$ and $\widehat{U}^{(2)}_N(u)$ are asymptotically $N^{1/2}\widehat{W}_1 e^u$ and $N^{1/2}\widehat{W}_2 e^u$ respectively. We write $\tilde{L}_{Nj}(u) = (1,0)$ if the j-th element of U_N is matched with a label associated with $\widehat{U}^{(1)}_N(u)$, and $(0,0)$ otherwise; similarly, $\tilde{L}_{Nj}(u) = (0,1)$ if matched with a label associated with $\widehat{U}^{(2)}_N(u)$ and $(0,0)$ otherwise. Then both K and K' belong to $\mathcal{S}_N(\tau_N + t_N(u))$ if $\sum_{j=1}^{|U_N|} \tilde{L}_{Nj}(u) = (0,0)$. The multivariate analogue of the Poisson approximation (2.30) (Roos, 1999, Theorem 1) gives

$$\begin{aligned} &\left| \mathbb{P}\Big[\sum_{j=1}^{|U_N|} \tilde{L}_{Nj}(u) \in \mathbb{N}^2 \,\Big|\, \mathcal{F}^{(u)}_N \cap \widetilde{A}_N \Big] - \exp\{-m_N(u, \widehat{W}_1) - m_N(u, \widehat{W}_2)\} \right| \\ &\qquad \leq\; c|U_N|^{-1}\{m_N(u, \widehat{W}_1) + m_N(u, \widehat{W}_2)\}, \end{aligned} \tag{2.32}$$

for a universal constant c. Hence, as before,

$$\begin{aligned} &\mathbb{P}[\{K, K'\} \subset \mathcal{S}_N(\tau_N + t_N(u)) \,|\, \mathcal{F}^+_{\tau_N} \cap \widetilde{A}_N] \\ &\qquad \sim \mathbb{E}\big\{\exp\{-m_N(u, \widehat{W}_1) - m_N(u, \widehat{W}_2)\} \,|\, \mathcal{F}^+_{\tau_N} \cap \widetilde{A}_N\big\} \;\sim\; \{\hat{s}(u)\}^2, \end{aligned} \tag{2.33}$$

by the independence of $\widehat{W}_1$ and $\widehat{W}_2$. But the joint probability can also be written as

$$\mathbb{E}\big\{\big(N^{-1}S_N(\tau_N+\lambda^{-1}\{\tfrac{1}{2}\log N+u\})\big)^2 \mid \mathcal{F}^+_{\tau_N}\cap\widetilde{A}_N\big\},$$

so that it follows from (2.31) and (2.33) that

$$\mathrm{Var}\left\{N^{-1}S_N(\tau_N+\lambda^{-1}\{\tfrac{1}{2}\log N+u\})\,|\,\mathcal{F}^+_{\tau_N}\cap\widetilde{A}_N\right\} \;\sim\; 0. \tag{2.34}$$

It now follows, by a standard argument, that, for any $\varepsilon>0$, conditional on $\mathcal{F}^+_{\tau_N}\cap\widetilde{A}_N$,

$$\mathbb{P}\Big[\sup_u |N^{-1}S_N(\tau_N+\lambda^{-1}\{\tfrac{1}{2}\log N+u\})-\hat{s}(u)|>\varepsilon\,\Big|\,\mathcal{F}^+_{\tau_N}\cap\widetilde{A}_N\Big] \;\to\; 0$$

as $N\to\infty$, and the theorem follows. □

Because of the factor λ^{-1} in the definition of $t_N(u)$, the quantity $\hat{s}(\lambda t)$ should match the solution $s(t)$ of (1.1). To see that this is so, note that, by considering the possibilities for the offspring of the first individual in $\widehat{Z}$, $\psi(\theta):=\mathbb{E}\{e^{-\theta\widehat{W}}\}$ satisfies the equation

$$\psi(\theta) \;=\; \exp\left\{-\mu\int_0^\infty (1-\psi(\theta e^{-\lambda w}))\,G(dw)\right\}. \tag{2.35}$$

Substituting $\theta=m_*e^{\lambda t}$, writing

$$s(t) \;=\; \hat{s}(\lambda t) \;=\; \psi(m_*e^{\lambda t}) \tag{2.36}$$

and taking logarithms recovers equation (1.2), with $\mu G(du)$ in place of $\beta(v)\,dv$. As for (1.2), equation (2.35) has many solutions, since, if $\psi(\theta)$ is a solution, so is $\psi_\alpha(\theta):=\psi(\alpha\theta)$, for any fixed $\alpha>0$. The condition $\psi(0)=1$, equivalent to $s(-\infty)=1$, is satisfied by all ψ_α. The relevant choice of solution to (2.35) is determined by matching $\mathbb{E}\widehat{W}$ with $-\psi'(0)$, or, in terms of (1.2), with $(m_*\lambda)^{-1}\lim_{t\to-\infty}e^{-\lambda t}(-Ds(t))$. A renewal equation for $\mathbb{E}\{\widehat{B}(t)e^{-\lambda t}\}$ gives the solution as

$$\mathbb{E}\widehat{W} \;=\; \lim_{t\to\infty}\mathbb{E}\{\widehat{B}(t)e^{-\lambda t}\} \;=\; \left\{\lambda\mu\int_0^\infty ve^{-\lambda v}\,G(dv)\right\}^{-1} \;=\; \frac{1}{m_*},$$

by the key renewal theorem. Thus Theorem 2.10 can be interpreted as a formal justification of the stochastic basis for the Kermack–McKendrick epidemic as described in Metz (1978), Section 4, under assumptions that are slightly more general, in that the point processes ξ are not required to be doubly stochastic, but are in some respects more restrictive as regards the choice of β. Since ψ is identified as the Laplace transform of a probability distribution, it is an analytic function in $\Re(\theta)>0$, which, with (2.36), proves Conjecture (f) in Metz (1978), p.120.

3 Refinements

3.1 Multitype epidemics

Very similar arguments can be carried through for epidemics in populations consisting of individuals of more than one type. Suppose that there are a finite number d of different types, with N_l individuals of type l, $1\le l\le d$, where $N_l\in\{\lfloor N\pi_l\rfloor,\lceil N\pi_l\rceil\}$, $\sum_{l=1}^d N_l=N$ and $\sum_{l=1}^d\pi_l=1$. Assume that type l individuals have independent and identically distributed point processes $\xi_i^{(l)}$, $1\le i\le N_l$, on $[d]\times\mathbb{R}_+$, with mean measures

$$\mathbb{E}\{\xi_1^{(l)}(k,du)\} \;=\; \mu_{lk}G_{lk}(du), \tag{3.1}$$

where $\int_0^\infty G_{lk}(du) = 1$ for all $1 \le k,l \le d$. Then an epidemic process can be constructed in the population, just as in the single type case, by beginning with a multitype branching process constructed from independent realizations of the $\xi_1^{(l)}$, $1 \le l \le d$, and then using random labelling within the members of each type to determine which transitions are to be retained in the epidemic process. The approximation arguments are very much as before. Asymptotically exponential growth and the analogues of (2.5) and (2.16), together with an asymptotically stable type distribution, hold in L_1 in the multitype setting. The asymptotic statements that we use in this section are all justified by Theorem 7.3 of Jagers (1989), who proves L_1 approximation for a wide variety of characteristics of the branching process in an even more general setting.

Remark 3.1. *It is perhaps more natural, especially when comparing the spread of the same epidemic in populations with different compositions of types, to assume a fixed value for the measures $\alpha_{lk}(du) := \mu_{lk}G_{lk}(du)/\pi_k$, rather than supposing that $\mu_{lk}G_{lk}$ remains the same for all N. The quantity $\alpha_{lk}(du)$ can be interpreted as representing the infection intensity measure of contacts with type k individuals made by a type l individual, in a population consisting entirely of individuals of type k. At least in Poisson process contact models, this would suggest taking $\mathbb{E}\{\xi_1^{(l,N)}(k,du)\} = \alpha_{lk}(du)N_k/N$ in a population of the composition given above, implying that $G_{lk}(du) = \alpha_{lk}(du)/\alpha_{lk}(\mathbb{R}_+)$ is fixed for all N, but that $\mu_{lk}^{(N)} = \alpha_{lk}(\mathbb{R}_+)N_k/N$ may vary with N. This differs from (3.1) inasmuch as N_k/N is not exactly equal to π_k. As in the single–type model, this minor difference entails no change in the theorems that we prove.*

We now assume that the matrix μ is irreducible, and that the distribution functions G_{lk} all satisfy Assumption 2; suppose also that the largest eigenvalue of μ is larger than 1, and write

$$\mu_{lk}(s) := \mu_{lk}\int_0^\infty e^{-su}\,G_{lk}(du).$$

Then the branching process has as Malthusian parameter the value $\lambda > 0$ for which $\mu(\lambda)$ has largest eigenvalue 1. We write ζ^T and η for the positive left and right eigenvectors of $\mu(\lambda)$ associated with eigenvalue 1, normalized such that $\zeta^T 1 = \zeta^T\eta = 1$. Let $B'(t) := (B'_l(t),\, 1 \le l \le d)$ denote the numbers of individuals of each type born up to time t. Then, if the branching process starts from a single individual of type i,

$$B'(t)e^{-\lambda t} \;\to\; W^{(i)}\zeta \text{ in } L_1 \text{ as } t \to \infty. \tag{3.2}$$

Here, $W^{(i)}$ is a random variable whose Laplace transform $\psi^{(i)}(s) := \mathbb{E}\{e^{-sW^{(i)}}\}$ satisfies the implicit equations

$$\psi^{(l)}(s) \;=\; \mathbb{E}\left\{\exp\left(\sum_{k=1}^d \int_0^\infty \log\psi^{(k)}(se^{-\lambda v})\xi^{(l)}(k,dv)\right)\right\}, \qquad 1 \le l \le d, \tag{3.3}$$

with $\mathbb{E}W^{(i)} = \eta_i/m_*^{(1)}$ and

$$m_*^{(1)} := \lambda\zeta^T(-D\mu(\lambda))\eta; \tag{3.4}$$

note that

$$(-D\mu(\lambda))_{lk} \;=\; \mu_{lk}\int_0^\infty ue^{-\lambda u}\,G_{lk}(du),$$

and that $m_*^{(1)}/\lambda$ is the multitype mean age at child bearing (Jagers (1989), p.195). Letting $V'_l(t)$ denote the set of times until birth of the unborn type l offspring of individuals born before t, it follows also that

$$e^{-\lambda t}|V'_l(t)| \;\to\; W^{(i)}c_l \text{ in } L_1, \tag{3.5}$$

with

$$\begin{aligned} c_l &:= \sum_{k=1}^{d} \zeta_k \mu_{kl} \int_0^\infty (1 - e^{-\lambda v})\, G_{kl}(dv) \\ &= \sum_{k=1}^{d} \zeta_k (\mu_{kl} - \mu_{kl}(\lambda)) \;=\; \sum_{k=1}^{d} \zeta_k \mu_{kl} - \zeta_l, \end{aligned} \tag{3.6}$$

and that, on $W^{(i)} > 0$,

$$\mathbb{E}^{(i)} \left(\sup_s \Big| |V_l'(t) \cap (s, \infty)|/|V_l'(t)| - (1 - F_l(s)) \Big| \right) \;\to\; 0, \tag{3.7}$$

where

$$1 - F_l(s) \;:=\; c_l^{-1} \sum_{k=1}^{d} \zeta_k \mu_{kl} \int_s^\infty (1 - e^{-\lambda(v-s)})\, G_{kl}(dv)\,, \tag{3.8}$$

replacing (2.5) and (2.6).

The backward branching process is similar, but has Poisson point processes $\hat{\xi}^{(l)}$ with intensity $\mu_{kl} G_{kl}(du)$ at $(k, u) \in [d] \times \mathbb{R}_+$. The matrix $\hat{\mu}(s)$ is given by $\mu(s)^T$, so that the Malthusian parameter is still λ, but the left and right eigenvectors at λ are swapped; the normalized versions are $\hat{\zeta}^T := \eta^T/H$ and $\hat{\eta} := H\zeta$, where $H := \sum_{k=1}^d \eta_k$. The backward random variables $\widehat{W}^{(l)} := \lim_{t\to\infty} e^{-\lambda t} \sum_{k=1}^d \widehat{B}_k(t)$ corresponding to the initial conditions $1 \le l \le d$ now have means $\hat{\eta}_l/m_*^{(1)} = H\zeta_i/m_*^{(1)}$, and their Laplace transforms $\hat{\psi}^{(l)}$ satisfy the equations

$$\hat{\psi}^{(l)}(s) \;=\; \exp\left(-\sum_{k=1}^{d} \mu_{kl} \int_0^\infty (1 - \hat{\psi}^{(k)}(s e^{-\lambda v}))\, G_{kl}(dv) \right), \qquad 1 \le l \le d. \tag{3.9}$$

As in (2.16), the empirical distribution of the ages at time t of l-individuals born before t also converges in L_1 to Exp(λ).

Now suppose that the forward branching process starts with a single type i individual. Define $\tau_N := \inf\{t > 0 \colon \sum_{l=1}^d B_l'(t) \ge \lfloor \sqrt{N} \rfloor\}$, so that $W^{(i)} e^{\tau_N} \sim \sqrt{N}$ as $N \to \infty$, from (3.2), and $|V_l'(\tau_N)| \sim c_l \sqrt{N}$, $1 \le l \le d$, from (3.5). Then run the backward branching process starting with a single type i' individual; at time $t_N(u) := \lambda^{-1}(\frac{1}{2} \log N + u)$, as in (2.7), we have $\widehat{B}(t_N(u)) \sim \sqrt{N} \widehat{W}^{(i')} e^{\lambda u} \hat{\zeta}$. Hence the mean number of pairs consisting of one element v of $V_l'(\tau_N)$ and one type l individual w born before $t_N(u)$ in the backward branching process, such that v is less than the age of w at $t_N(u)$, is asymptotically given by

$$\begin{aligned} &\{c_l \sqrt{N}\}\, \{\sqrt{N} \widehat{W}^{(i')} e^{\lambda u} \hat{\zeta}_l\} \int_0^\infty \lambda e^{-\lambda s} F_l(s)\, ds \\ &\qquad = N \widehat{W}^{(i')} e^{\lambda u} \hat{\zeta}_l \int_0^\infty \lambda v e^{-\lambda v} \sum_{k=1}^{d} \mu_{kl} \zeta_k\, G_{kl}(dv). \end{aligned}$$

Thus, when the individuals corresponding to the $V_l'(\tau_N)$ and the type l individuals in the backward branching process are randomly labelled in constructing the epidemic process, the mean number of such pairs that have the same labels is asymptotically given by

$$\pi_l^{-1} \widehat{W}^{(i')} e^{\lambda u} \hat{\zeta}_l \int_0^\infty \lambda v e^{-\lambda v} \sum_{k=1}^{d} \mu_{kl} \zeta_k\, G_{kl}(dv),$$

and hence the probability that there is no such pair of any type l, $1 \le l \le d$, is asymptotically given by $\exp\{-\widehat{W}^{(i')}e^{\lambda u}m_*^{(2)}\}$, where

$$m_*^{(2)} \;:=\; \int_0^\infty \lambda v e^{-\lambda v} \sum_{k=1}^{d}\sum_{l=1}^{d} \zeta_k\, \alpha_{kl}(dv)\eta_l/H. \tag{3.10}$$

Arguing as in the case of a single type, we have the following theorem, in which $\mathcal{F}^+_{\tau_N}$ denotes the precise analogue of the σ-algebra having the same name in the single type case, and $S_{Nl}(t)$ is the number of type l susceptibles at time t.

Theorem 3.2. *Suppose that the multitype forward branching process is supercritical and has offspring distributions with finite second moments; suppose also that Assumption 2 holds for each G_{lk}. Then there exists an event $\widetilde{A}_N \in \mathcal{F}^+_{\tau_N}$ such that $\mathbb{P}[\widetilde{A}_N^c] \to 0$ as $N \to \infty$, for which*

$$\mathbb{P}\Big[\sup_u |(Np_l)^{-1}S_{Nl}(\tau_N + \lambda^{-1}\{\tfrac{1}{2}\log N + u\}) - \hat{s}_l(u)| > \varepsilon \,\Big|\, \mathcal{F}^+_{\tau_N} \cap \widetilde{A}_N \cap \{\tau_N < \infty\}\Big] \;\to\; 0$$

as $N \to \infty$, where $\hat{s}_l$ is the decreasing function given by

$$\hat{s}_l(u) \;:=\; \hat{\psi}^{(l)}(e^u m_*^{(2)}),$$

where the $\hat{\psi}^{(l)}$ satisfy (3.9) with $-D\hat{\psi}^{(l)}(0) = H\zeta_l/m_*^{(1)}$, and where $m_*^{(1)}$ is defined in (3.4) and $m_*^{(2)}$ in (3.10).

3.2 A configuration model

In this section, we consider a different model of epidemic spread. In those considered so far, an infected individual chooses to infect a number of randomly chosen individuals, and the individuals chosen are not taken into account in this choice. Now we suppose that pairs of individuals are either acquainted with one another or are not, so that acquaintanceship determines a graph on the set of individuals, and we assume that infectious contacts can only be made between graph neighbours. This yields a more symmetric description of the contact process, and, as a result, the forward and backward branching approximations can be expected to look more similar. We shall, for simplicity, assume that there is a finite, N-independent upper bound K on the number of acquaintances that an individual may have; note that this immediately rules out any Poisson distribution of offspring in an approximating branching process, so that the backward branching processes from such a model have to be different from those in the previous sections.

To make further progress, we assume that the acquaintanceship graph is nonetheless rather randomly constituted within the population, according to the following construction. We assume that N_k members of the population are 'type k' individuals, who have exactly k acquaintances, with $\sum_{k=1}^K N_k = N$ and $N_k \in \{\lfloor N\pi_k \rfloor, \lceil N\pi_k \rceil\}$, for fixed $\pi_1, \ldots, \pi_K$, and with $M := \sum_{k=1}^K kN_k$ even. Think of a type k individual as having k half–edges, and join the half–edges into edges by means of a random matching of the M half–edges, determining the acquaintanceship graph. This graph may have some loops and multiple edges, but they are few, and we shall ignore their effects. Thus the method of assigning which individuals are acquaintances remains essentially random, but the propensities of each individual are respected when determining whether they are acquainted or not. We then assume that an infected type k individual makes contact with a given type l acquaintance at a random time after infection that has (possibly defective) distribution function G_{kl} and is independent of all other contact times; we suppose also that a type k individual remains infectious for a random time with (possibly defective)

distribution Φ_k, again independently of everything else. If we specialize to the case where the distributions G_{kl} are all identical and equal to Exp(α), and that the Φ_k are all identical and equal to Exp(β), then the model of Volz (2008) (in the case of a finite number K of possible contact numbers) is recovered. If all infections are assumed to be of infinite duration, so that $\Phi_k(t) = 0$ for all t, and if the G_{kl} are all identical, then the first passage percolation model considered by Bhamidi, van der Hofstad & Hooghiemstra (2012) is obtained.

As in the previous models, the key effort lies in determining the probability that an initially chosen individual infects another randomly chosen individual before a specified time t. To do so, construct the association graph by starting from the initial individual as root vertex, and matching its half–edges by random choice from the set of all half–edges; then attach the infectious period to the initial individual, and the lengths of time to potentially infectious contact to the edges. This yields a set of infected vertices, together with the times of their infection, some of which may be infinite. Now continue by matching the remaining half–edges associated with the first of these vertices (if any) to be infected, attaching the infectious period to the chosen vertex, and adding the lengths of time to potentially infectious contact (infinite, if longer than the infectious period) for each edge to the time of infection of the chosen vertex, so as to yield the times of infection of newly infected vertices; this augments the set of infected vertices. Proceed in this way, always choosing for development the infected vertex with unmatched half–edges that has the smallest time of infection, until the first time that either at least $\lfloor\sqrt{N}\rfloor$ vertices have been infected or the infection dies out. In the former case, there remains a set of infected vertices whose subsequent contact history has not been explored. If a half–edge is picked for a second or subsequent time, ignore the choice and re-sample until a new one is chosen; if a vertex is chosen that has already been infected, ignore it for future development. As in the previous arguments, for the lengths of time in which we are interested, there are a few such repeated samples, but few enough that they can be ignored.

For the susceptibility graph seen backwards from a randomly chosen individual, carry out essentially the same procedure for a specified time; the only difference is the vertex to which the infectious period is attached, being that of the child, rather than the parent. Half–edges that have previously been used, including those that were used in the forward process, are discarded and re-sampled; the half–edges that are associated with the set of infected but unexplored vertices from the forward phase are still available for choice, and are those that close chains of infection.

If repeats are ignored, the infection process as seen from the initial individual becomes a branching process with K types. In the branching process, a type k individual (other than the initial individual) has $k-1$ offspring, corresponding to the $k-1$ half–edges that remain to be connected after a type k individual has been encountered in growing the association graph, and each of these is of type l with probability lp_l/m, where $m = \sum_{l'=1}^{K} l' p_{l'}$, chosen from the size–biased transform of the frequency distribution $(p'_l,\, 1 \le l' \le K)$. As before, the difference between the process with this distribution and that with offspring probabilities lN_l/M is negligible for our purposes. The type k individual also has an infectious period randomly assigned to it from the distribution Φ_k, and the times to contact along the different edges are assigned independently from the appropriate distributions G_{kl}. This yields an age–dependent multi-type branching process, in which times to birth may be infinite (if the sampled time to contact is itself infinite, or exceeds the infectious period of the parent), and the times of birth of the descendants of a given individual are dependent, because they are finite only if they do not exceed the infectious period of the common parent.

Seen from the randomly chosen individual, the backward branching process is very

much the same. The offspring distribution is identical, but the infection times of the offspring of a given individual, although having the same marginal distributions as before, are now independent, because the relevant infectious period, determining whether a contact results in infection, is that of the child, and not of the parent. Because the basis of the construction is the fixed set of half–edges, the problems that arose in Section 2.2, because the offspring distribution of the backward branching process was not fixed for all N, no longer appear (except for the trivial differences between lp_l/m and lN_l/M); more importantly, choosing the contact times for type k – type l contacts *independently* from G_{kl} and the infectious periods independently from the Φ_k means that the times to birth in the backward branching process have distributions that do not depend on N, so that there is no need for an analogue of Corollary 2.5, and hence the G_{kl} need not satisfy the tail condition in Assumption 2. Of course, the numbers of offspring of the different types are bounded, so that the corresponding moment conditions are automatically satisfied.

The argument now proceeds much as for the multitype process of the previous section. Once again, the asymptotic statements for the branching processes are justified by Jagers (1989), Theorem 7.3. The matrix μ is defined analogously by

$$\mu_{lk}(s) \;:=\; (l-1)\{kp_k/m\} \int_0^\infty e^{-su}\,(1-\Phi_l(u))G_{lk}(du) \;=:\; (l-1)\{kp_k/m\}U_{lk}(s),$$

say, and we write $\mu_{lk} := \mu_{lk}(0)$; note that μ_{lk} need no longer be the expected number of offspring, since $U_{lk}(0)$ is typically less than 1. Because of the factor $(l-1)$, $\mu(s)$ is reducible. Supposing that all the p_k and all the $U_{lk}(s)$ are positive, we can write the irreducible non-negative matrix $\mu^{(1)}(s)$, obtained from $\mu(s)$ by removing the first row and column, as $D_1U^{(1)}(s)D_2$, where $D_1 := \mathrm{diag}(1,2,\ldots,K-1)$ and $D_2 := m^{-1}\mathrm{diag}(2p_2,\ldots,Kp_K)$. Assume that the matrix $\mu^{(1)}(0)$ has dominant eigenvalue larger than 1, and define the Malthusian parameter λ to be such that $\mu^{(1)}(\lambda)$ has dominant eigenvalue equal to 1; let ${\zeta^{(1)}}^T$ and $\eta^{(1)}$ be associated left and right eigenvectors. Then the left and right eigenvectors of $\mu^{(1)}(\lambda)$ with eigenvalue 1 are given by $\zeta^T := Z^{-1}({\zeta^{(1)}}^T\mu(\lambda)\varepsilon^{(1)}, {\zeta^{(1)}}^T)$ and $\eta := H^{-1}(0,{\eta^{(1)}}^T)^T$, where Z and H are chosen so that $\zeta^T 1 = \zeta^T\eta = 1$; here, $\varepsilon^{(1)}$ denotes the first coordinate vector.

Let $B'(t) := (B'_l(t),\, 1\le l\le d)$ denote the numbers of individuals of each type born up to time t; then

$$B'(t)e^{-\lambda t} \;\to\; W_*^{(i)}\zeta \text{ in } L_1 \tag{3.11}$$

as $t\to\infty$, if the initial individual has type i. The distribution of $W_*^{(i)}$ is not quite the one that would be expected when starting the branching process with a typical type i individual, because the *initial* type i individual has i offspring, instead of $i-1$. However, it can easily be deduced from the Laplace transforms $(\psi^{(l)}(s),\, 1\le l\le K)$ of the limiting random variables for the branching process that has all individuals, including the initial one, obeying the same rules. These solve a system of implicit equations that can be deduced from (3.3). Here, the quantity within the expectation in (3.3) can be written as

$$\prod_{r=1}^{l-1}\left\{\left(\psi^{(K_r)}(se^{-\lambda V_r})\right)^{I[V_r\le T]}\right\},$$

where T denotes the infectious period of the type l individual, and K_r denotes the type and V_r the contact time of the r-th of his $(l-1)$ acquaintances. T and $(K_r,\, 1\le r\le l-1)$ are independent, and, given $K_r=k$, V_r is drawn independently of everything else from

the distribution G_{lk}. Thus (3.3) reduces here to the system

$$\psi^{(l)}(s) \;=\; \int_{[0,\infty]} \left\{ m^{-1} \sum_{k=1}^{K} k p_k \left(\int_{[0,t]} \psi^{(k)}(se^{-\lambda v}) G_{lk}(dv) + [1 - G_{lk}(t)] \right) \right\}^{l-1} \Phi_l(dt), \tag{3.12}$$

for $1 \le l \le d$, with $-(D\psi^{(l)})(0) = \eta_l / m_*^{(1)}$ and

$$m_*^{(1)} \;:=\; \lambda \zeta^T (-D\mu)(\lambda) \eta; \tag{3.13}$$

the Laplace transform of the distribution of $W_*^{(l)}$ is then given by

$$\mathbb{E}\left\{ e^{-sW_*^{(l)}} \right\} \;=\; \int_{[0,\infty]} \left\{ m^{-1} \sum_{k=1}^{K} k p_k \left(\int_{[0,t]} \psi^{(k)}(se^{-\lambda v}) G_{lk}(dv) + [1 - G_{lk}(t)] \right) \right\}^{l} \Phi_l(dt), \tag{3.14}$$

for $1 \le l \le d$. Here, we have

$$(-D\mu)(\lambda)_{lk} \;=\; (l-1)\{k p_k / m\} \int_0^\infty u e^{-\lambda u} \, (1 - \Phi_l(u)) G_{lk}(du).$$

Letting $V_l'(t)$ denote the set of times until birth of the unborn type l offspring of individuals born before t, it follows also that, if the initial individual is of type i, then

$$e^{-\lambda t} |V_l'(t)| \;\to\; W_*^{(i)} c_l \quad \text{in } L_1, \tag{3.15}$$

with

$$\begin{aligned} c_l \;&:=\; \sum_{k=1}^{K} \zeta_k (k-1)\{l p_l / m\} \int_0^\infty (1 - e^{-\lambda v}) \, (1 - \Phi_k(v)) G_{kl}(dv) \\ &=\; \sum_{k=1}^{K} \zeta_k (\mu_{kl}(0) - \mu_{kl}(\lambda)) \;=\; \sum_{k=1}^{K} \zeta_k \mu_{kl} - \zeta_l. \end{aligned} \tag{3.16}$$

Furthermore, on $W_*^{(i)} > 0$, as for (3.7) and (3.8),

$$\mathbb{E}^{(i)} \left(\sup_s \left| |V_l'(t) \cap (s, \infty)| / |V_l'(t)| - (1 - F_l(s)) \right| \right) \;\to\; 0, \tag{3.17}$$

where

$$1 - F_l(s) \;:=\; c_l^{-1} \sum_{k=1}^{K} \zeta_k (k-1)\{l p_l / m\} \int_s^\infty (1 - e^{-\lambda(v-s)}) \, (1 - \Phi_k(v)) G_{kl}(dv) \, . \tag{3.18}$$

The backward branching process is similar; we now have

$$\hat{\mu}_{lk}(s) \;:=\; (l-1)\{k p_k / m\} U_{kl}(s),$$

once again reducible, with $\hat{\mu}^{(1)}(s) = D_1 U^{(1)^T}(s) D_2$ irreducible. It can be checked that the Malthusian parameter is still λ. The matrix $\hat{\mu}^{(1)}(\lambda)$ has left and right eigenvectors $\hat{\zeta}^{(1)^T} = \eta^{(1)^T} D_2 D_1^{-1}$ and $\hat{\eta}^{(1)} = D_2^{-1} D_1 \zeta^{(1)}$ with eigenvalue 1, and the corresponding left and right eigenvectors of $\hat{\mu}(\lambda)$ are given by $\hat{\zeta}^T = \widehat{Z}^{-1}(\eta^{(1)^T} D_2 U^T(\lambda) D_2 \varepsilon^{(1)}, \hat{\zeta}^{(1)^T})$ and $\hat{\eta} = \widehat{H}^{-1}(0, \hat{\eta}^{(1)^T})^T$, where $\widehat{Z}$ and $\widehat{H}$ are chosen to make $\hat{\zeta}^T 1 = \hat{\zeta}^T \hat{\eta} = 1$; in particular, it follows that $\widehat{Z}\widehat{H} = 1$, and that the value of $m_*^{(1)}$ deduced from (3.13) for the backward process is the same as $m_*^{(1)}$. The limiting random variable $\widehat{W}_*^{(i)}$ for the backward process starting with a single individual of type i, satisfying

$$\widehat{B}'(t) e^{-\lambda t} \;\to\; \widehat{W}_*^{(i)} \hat{\zeta} \text{ in } L_1 \, , \tag{3.19}$$

once again has a distribution whose Laplace transform $\hat{\psi}_*^{(i)}$ can be found from the solutions to a set of implicit equations belonging to the backward branching process whose individuals, including the initial individual, all follow the same rules. This branching process has offspring that behave independently of one another as regards both type and time of birth, so that, denoting the Laplace transforms of the limit random variables with the different initial conditions by $(\hat{\psi}^{(l)}, 1 \le l \le K)$, we have $\hat{\psi}^{(l)}(s) = \{\hat{\psi}_0^{(l)}(s)\}^{l-1}$, where the $\hat{\psi}_0^{(l)}$ satisfy the equations

$$\begin{aligned}\hat{\psi}_0^{(l)}(s) &= m^{-1}\sum_{k=1}^{K} kp_k \left\{ \int_{(0,\infty)} \{\hat{\psi}_0^{(k)}(se^{-\lambda v})\}^{k-1}\,(1-\Phi_k(v))G_{kl}(dv) + (1-U_{kl}(0)) \right\} \\ &= 1 - m^{-1}\sum_{k=1}^{K} kp_k \int_{(0,\infty)} (1-\{\hat{\psi}_0^{(k)}(se^{-\lambda v})\}^{k-1})(1-\Phi_k(v))G_{kl}(dv), \qquad (3.20)\end{aligned}$$

for $1 \le l \le d$. Since $(-D\hat{\psi}^{(l)})(0) = \hat{\eta}_l/m_*^{(1)}$, the side condition for solving (3.20) is $(-D\hat{\psi}_0^{(l)})(0) = \hat{\eta}_l/\{(l-1)m_*^{(1)}\} = \widehat{H}^{-1}\{m/lp_l\}\zeta_l$, $l \ge 2$, with $\hat{\psi}_0^{(1)}(s) = 1$ for all s. The Laplace transform $\hat{\psi}_*^{(i)}$ of $\widehat{W}_*^{(i)}$ is then given by $\{\hat{\psi}_0^{(i)}\}^i$. As in (2.16), the empirical distribution of the ages at time t of l-individuals born before t also converges to Exp(λ).

Now suppose that the forward branching process starts with a single type i individual (having i offspring). Define $\tau_N := \inf\{t > 0\colon \sum_{l=1}^{K} B_l'(t) \ge \lfloor\sqrt{N}\rfloor\}$, so that $W_*^{(i)}e^{\tau_N} \sim \sqrt{N}$ as $N \to \infty$, from (3.11), and $|V_l'(\tau_N)| \sim c_l\sqrt{N}$, $1 \le l \le K$, from (3.15). Then run the backward branching process starting with a single type i' individual; at time $t_N(u) := \lambda^{-1}(\frac{1}{2}\log N + u)$, we have $\widehat{B}(t_N(u)) \sim \sqrt{N}\widehat{W}_*^{(i')}e^{\lambda u}\hat{\zeta}$. Hence the mean number of pairs of individuals consisting of an element v of $V_l'(\tau_N)$ and a type l individual w born before $t_N(u)$ in the backward branching process, such that v is less than the age of w at $t_N(u)$, is asymptotically given by

$$\begin{aligned}&\{c_l\sqrt{N}\}\,\{\sqrt{N}\widehat{W}_*^{(i')}e^{\lambda u}\hat{\zeta}_l\}\int_0^\infty \lambda e^{-\lambda s}F_l(s)\,ds \\ &\qquad = N\widehat{W}_*^{(i')}e^{\lambda u}\hat{\zeta}_l \int_0^\infty \lambda v e^{-\lambda v}\sum_{k=1}^{K}\zeta_k(k-1)\{lp_l/m\}\,(1-\Phi_k(v))G_{kl}(dv).\end{aligned}$$

Any such pair is realized as identical individuals in the epidemic process with asymptotic probability $(l-1)/Nlp_l$, since the element v has only $(l-1)$ half edges available to be matched, out of a total number of half–edges from type l individuals that is still asymptotically Nlp_l. Thus the mean number of such pairs that correspond to actual matches is asymptotically given by

$$\frac{(l-1)}{lp_l}\,\widehat{W}_*^{(i')}e^{\lambda u}\hat{\zeta}_l \int_0^\infty \lambda v e^{-\lambda v}\sum_{k=1}^{K}\zeta_k(k-1)\{lp_l/m\}\,(1-\Phi_k(v))\,G_{kl}(dv),$$

and hence the probability that there is no such pair of any type l, $1 \le l \le d$, is asymptotically given by $\exp\{-\widehat{W}_*^{(i')}e^{\lambda u}m_*^{(2)}\}$, where

$$m_*^{(2)} := \frac{1}{m}\sum_{k=1}^{K}\sum_{l=1}^{K}\zeta_k(k-1)(l-1)\hat{\zeta}_l\int_0^\infty \lambda v e^{-\lambda v}(1-\Phi_k(v))G_{kl}(dv)\,. \qquad (3.21)$$

These assertions, and the analogous assertions about the probability of two randomly chosen individuals being infected by the initial individual, can be proved by the methods introduced in Section 2, and lead to the following theorem. Here, $\mathcal{F}_{\tau_N}^+$ denotes the σ-algebra associated with the (forward) infection process until $\lfloor\sqrt{N}\rfloor$ infections have occurred, and $S_{Nl}(t)$ denotes the number of type l susceptibles at time t.

Theorem 3.3. *Suppose that the forward branching process is supercritical. Then there exists an event $\widetilde{A}_N \in \mathcal{F}^+_{\tau_N}$ such that $\mathbb{P}[\widetilde{A}_N^c] \to 0$ as $N \to \infty$, for which*

$$\mathbb{P}\Big[\sup_u |(Np_l)^{-1} S_{Nl}(\tau_N + \lambda^{-1}\{\tfrac{1}{2}\log N + u\}) - \hat{s}_l(u)| > \varepsilon \,\Big|\, \mathcal{F}^+_{\tau_N} \cap \widetilde{A}_N \cap \{\tau_N < \infty\}\Big] \;\to\; 0$$

as $N \to \infty$, where $\hat{s}_l$ is the decreasing function given by

$$\hat{s}_l(u) \;:=\; \hat{\psi}_*^{(l)}(e^u m_*^{(2)}),$$

where $\hat{\psi}_*^{(l)}$ and $m_*^{(2)}$ are as defined above. In particular, the total proportion of susceptibles $N^{-1}\sum_{l=1}^K S_{Nl}(\tau_N + \lambda^{-1}\{\frac{1}{2}\log N + u\})$ is well approximated by $\sum_{l=1}^K p_l \hat{s}_l(u)$, uniformly in u.

The general formulation above simplifies, if the distributions Φ_k of infectious period and G_{kl} of contact times are the same for all choices of the indices. In this case, the matrix $\mu(s)$ is of rank one and is given by

$$\mu_{lk}(s) \;:=\; (l-1)\{kp_k/m\} \int_0^\infty e^{-su}\,(1-\Phi(u))G(du) \;=:\; U(s)(l-1)\{kp_k/m\}.$$

The positive eigenvalue is $U(s)m_{(2)}/m$, where $m_{(2)} := \sum_{k=1}^K k(k-1)p_k$, the process is supercritical if $m_{(2)}/m > 1/U(0)$, where $U(0) = \int_{(0,\infty)}(1-\Phi(u))G(du)$, and λ is such that $U(\lambda) = m/m_{(2)}$. The eigenvectors for the forward and backward processes are equal, with $\zeta_i = \hat{\zeta}_i = ip_i/m$ and $\eta_i = \hat{\eta}_i = (i-1)m/m_{(2)}$. The quantities $m_*^{(1)}$ and $m_*^{(2)}$ become

$$m_*^{(1)} \;=\; m_0 m_{(2)}/m \qquad \text{and} \qquad m_*^{(2)} \;=\; m_0 m_{(2)}^2/m^3,$$

where $m_0 := \int_0^\infty \lambda v e^{-\lambda v}(1-\Phi(v))G(dv)$. The equations (3.20) can be much more neatly expressed, because the functions $\hat{\psi}_0^{(l)}$ are now the same for all l, reflecting that the backward process of half–edges is equivalent to a single–type branching process. They reduce to the single equation

$$\begin{aligned}\hat{\psi}_0(s) \;&=\; m^{-1}\sum_{k=1}^K kp_k \left\{\int_{(0,\infty)} \{\hat{\psi}_0(se^{-\lambda v})\}^{k-1}\,(1-\Phi(v))G(dv) + (1-U(0))\right\} \\ &=\; 1 - \int_{(0,\infty)} \{1 - m^{-1}g'(\hat{\psi}_0(se^{-\lambda v}))\}(1-\Phi(v))G(dv), \qquad (3.22)\end{aligned}$$

where $g(s) := \sum_{k=1}^K p_k s^k$, and the initial condition is $(-D\hat{\psi}_0)(0) = \hat{\eta}_l/\{(l-1)m_*^{(1)}\} = m/\{m_{(2)}m_*^{(1)}\}$. To express $\hat{s}_l(u) = \{\hat{\psi}_0(e^{\lambda u}m_*^{(2)})\}^l$ more concisely, we write $h_s(u) := \hat{\psi}_0(se^{\lambda u})$; then (3.22) implies that $h := h_s$ satisfies the equation

$$h(u) \;=\; 1 - \int_{(0,\infty)} \{1 - m^{-1}g'(h(u-v))\}(1-\Phi(v))G(dv), \qquad (3.23)$$

with initial condition $\lim_{u\to-\infty}\{e^{-\lambda u}(Dh)(u)\} = \lambda s$; so $\hat{s}_1(u) = h_{m_*^{(2)}}(u)$ satisfies (3.23) with $\lim_{u\to-\infty}\{e^{-\lambda u}(D\hat{s}_1)(u)\} = \lambda m_*^{(2)}$, and $\hat{s}_l = (\hat{s}_1)^l$.

In the case of the Volz (2008) model, there is further simplification, because of the explicit forms $\Phi(v) = 1 - e^{-\beta v}$ and $G(dv) = \alpha e^{-\alpha v}dv$. In this case, a deterministic law of large numbers starting with an asymptotically positive initial proportion of infectious individuals was established by Decreusefond *et al.* (2012). With such an initial condition, the randomness inherent in the initial stages of development, reflected in the presence of τ_N in the statement of Theorem 3.3, plays no significant part. In the Volz

setting, explicit formulae for $\lambda = \alpha m_{(2)}/m - \beta$ and $m_0 = \lambda\alpha(\lambda+\alpha+\beta)^{-2}$ can be written down, and equation (3.23) can be expressed as

$$\begin{aligned} h(u) &= 1 - \int_{(0,\infty)} \{1 - m^{-1}g'(h(u-v))\}\alpha e^{-(\alpha+\beta)v}\, dv \\ &= \frac{\beta}{\alpha+\beta} + \frac{1}{m}\int_{-\infty}^{u} g'(h(w))\alpha e^{-(\alpha+\beta)(u-w)}\, dw\,. \end{aligned}$$

Differentiation with respect to u then yields the following autonomous differential equation for $h = h(t)$:

$$\frac{dh}{dt} \;=\; \frac{\alpha}{m}g'(h) - (\alpha+\beta)h + \beta \;=\; (\alpha+\beta)(\tilde{f}(h) - h), \tag{3.24}$$

where $\tilde{f}(s)$ is the probability generating function $(\alpha g'(s)/m + \beta)/(\alpha+\beta)$. In particular, it follows that $h(\infty)$ is the solution $\tilde{q}$ smaller than 1 to the equation $\tilde{f}(s) = s$, and hence that the asymptotic final proportion of susceptible individuals at the end of a large outbreak is given by $g(\tilde{q})$.

Remark 3.4. *Volz (2008) expresses the equations for the development of the epidemic as the solutions to a system of three coupled differential equations for the variables h, p_I and p_S:*

$$\begin{aligned} \frac{dh}{dt} &= -\alpha p_I h; \qquad \frac{dp_S}{dt} \;=\; \alpha p_S p_I \left(1 - \frac{hg''(h)}{g'(h)}\right); \\ \frac{dp_I}{dt} &= \alpha p_S p_I \frac{hg''(h)}{g'(h)} - \alpha p_I(1-p_I) - \beta p_I. \end{aligned}$$

The equivalence of these equations to (3.24), together with $p_I = 1 - \frac{g'(h)}{mh} + \frac{\beta}{\alpha}\{1 - \frac{1}{h}\}$ and $p_S = \frac{g'(h)}{mh}$, was noted by Miller (2011).

References

[1] Aldous, D. J.: When knowing early matters: gossip, percolation and Nash equilibria. In: Prokhorov and Contemporary Probability Theory, Eds. A. N. Shiryaev, S. R. S. Varadhan & E. L. Presman. *Springer*, 2012.

[2] Ball, F. G.: The threshold behaviour of epidemic models. *J. Appl. Probab.* **20**, (1983), 227–241. MR-0698527

[3] Ball, F. G. and Donnelly, P.: Strong approximations for epidemic models. *Stoch. Procs Applics.* **55**, (1995), 1–21. MR-1312145

[4] Barbour, A. D., Holst, L. and Janson, S.: Poisson approximation. *Clarendon Press*, Oxford, 1992. MR-1163825

[5] Barbour, A. D. and Reinert, G.: Asymptotic behaviour of gossip processes and small world networks. *J. Appl. Probab.* (to appear), arXiv:1202.5895.

[6] Bhamidi, S., van der Hofstad, R. and G. Hooghiemstra, G.: Universality for first passage percolation on sparse random graphs, arXiv:1210.6839.

[7] Brauer, F.: The Kermack-McKendrick epidemic model revisited. *Mathematical Biosciences* **198**, (2005), 119–131. MR-2187870

[8] Brauer, F. and Castillo–Chávez, C. Mathematical Models in Population Biology and Epidemiology, 2^{nd} ed., *Springer*, New York, 2012. MR-3024808

[9] Decreusefond, L., Dhersin, J.-S., Moyal, P. and Tran, V. C.: Large graph limit for an SIR process in random network with heterogeneous connectivity. *Ann. Appl. Probab.* **22**, (2012), 541–575. MR-2953563

[10] Diekmann, O.: Limiting behavious in an epidemic model. *Nonlinear analysis; Theory Meth. Applics* **1**, (1977), 459–470. MR-0624451

[11] Diekmann, O. and Heesterbeek, J. A. P.: Mathematical epidemiology of infectious diseases. *Wiley*, Chichester, 2000. MR-1882991

[12] Gupta, S.D., Lal, V., Jain, R. and Gupta, O. P.: Modeling of H1N1 Outbreak in Rajasthan: Methods and Approaches. *Indian J. Community Med.* **36**, (2011), 36–38.

[13] Jagers, P.: Branching Processes with Biological Applications. *Wiley*, London, 1975. MR-0488341

[14] Jagers, P.: General branching processes as Markov random fields. *Stoch. Procs Applics* **32**, (1989), 183–212. MR-1014449

[15] Kendall, D. G.: Deterministic and stochastic epidemics in closed populations. *Proc. Third Berkeley Symp. Math. Statist. Probab.* **4**, (1956), 149–165. *University of California Press*, Berkeley. MR-0084936

[16] Kermack, W. O. and McKendrick, A. G.: Contributions to the mathematical theory of epidemics, part I. *Proc. Roy. Soc. Edin. A* **115**, (1927), 700–721; *Bull. Math. Biol.* **53**, (1991), 33–55.

[17] Massart, P.: The tight constant in the Dvoretzky–Kiefer–Wolfowitz inequality. *Ann. Probab.* **18**, (1990), 1269–1283. MR-1062069

[18] McDiarmid, C.: Concentration. In: Probabilistic Methods for Algorithmic Discrete Mathematics, M. Habib, C. McDiarmid, J. Ramirez-Alfonsin and B. Reed (eds.), *Springer*, Berlin, 195–248, 1998. MR-1678578

[19] Metz, J. A. J.: The epidemic in a closed population with all susceptibles equally vulnerable; some results for large susceptible populations and small initial infections. *Acta Biotheor.* **27**, (1978), 75–123.

[20] Miller, J. C.: A note on a paper by Erik Volz: SIR dynamics in random networks. *J. Math. Biol.* **62**, (2011), 349–358. MR-2771177

[21] Moore, C. and Newman, M. E. J.: Epidemics and percolation in small-world networks. *Phys. Rev. E* **61**, (2000), 5678–5682.

[22] Nerman, O.: On the convergence of the supercritical general (C-M-J) branching process. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **57**, (1981), 365–395. MR-0629532

[23] Roos, B.: On the rate of multivariate Poisson convergence. *J. Multiv. Analysis* **69**, (1999), 120–134. MR-1701409

[24] Volz, E.: SIR dynamics in random networks with heterogeneous connectivity. *J. Math. Biol.* **56**, (2008), 293–310. MR-2358436

[25] Whittle, P.: The outcome of a stochastic epidemic – a note on Bailey's paper. *Biometrika* **42**, (1955), 116–122. MR-0070099

Acknowledgments. This work was carried out in part while ADB was Saw Swee Hock Professor of Statistics at the National University of Singapore. ADB was supported in part by Australian Research Council Grants Nos DP120102728 and DP120102398. We are grateful to Peter Jagers, Hans Heesterbeek and Odo Diekmann for a number of helpful discussions, and to the referees for their pertinent comments. ADB thanks the mathematics departments of the University of Melbourne, Monash University and the University of Queensland, and the Department of Statistics and Applied Probability and the Institute for Mathematical Sciences at the National University of Singapore, for their kind hospitality while part of the work was undertaken. GDR thanks EPSRC, the BBSRC and the Oxford Martin School for their support.