

# Couplings for locally branching epidemic processes

A. D. Barbour\*

Universität Zürich and National University of Singapore

## Abstract

The asymptotic behaviour of many locally branching epidemic models can, at least to first order, be deduced from the limit theory of two branching processes. The first is the ‘Whittle’ branching approximation to the early stages of the epidemic, the phase in which approximately exponential growth takes place. The second is the ‘susceptibility’, or backward branching process, that approximates the contact history that would lead to an individual becoming infected. The simplest coupling arguments for demonstrating the closeness of the branching process approximations do not keep the processes identical for quite long enough, so that arguments showing that the differences are unimportant are also needed. In this paper, we show that, in some models, couplings can be constructed that are sufficiently accurate for this extra step to be dispensed with.

**Keywords:** Coupling, epidemic process, branching process approximation, deterministic approximation

**AMS Classification:** 92H30; 60K35, 60J85

## 1 Introduction

Stochastic epidemic modelling is one of the classical fields of applied probability. Originally, papers on the subject appeared in a wide variety of journals. McKendrick’s 1925 paper in the Proceedings of the Royal Society of Edinburgh and Greenwood’s 1931 paper in the Journal of Hygiene were early instances, and the seminal papers of Bartlett (1949) and Bailey (1953) appeared in the Journal of the Royal Statistical Society B and Biometrika, respectively. However, since the Journal of Applied Probability was founded in 1964, such articles have had a natural home. Two notable examples are Sellke’s (1983) paper, in which an ingenious new construction of the Markovian SIR epidemic model was introduced, and Ball’s (1983) coupling, in which a sequence of epidemic processes and a branching process are constructed together on the same probability space, in such a way that their paths are all identical during their early development.

---

\*Institut für Mathematik, Universität Zürich, Winterthurertrasse 190, CH-8057 ZÜRICH. Work completed while the author was Saw Swee Hock Professor of Statistics at the National University of Singapore, and supported in part by Australian Research Council Grants Nos DP120102728 and DP120102398.

The topic of this paper is directly concerned with two classical results that have strongly influenced the development of epidemic modelling. One is Whittle's (1955) birth and death approximation to the early stages of the Markovian SIR epidemic model, the precursor to Ball's coupling idea, and the second is Kendall's (1956) recognition of the solution to Kermack and McKendrick's (1927) differential equations as an approximation to the paths of the Markovian SIR epidemic, once the initial phase has come to an end. In a recent paper, Chatterjee & Durrett (2011) studied the Aldous (2012) gossip process, which can be interpreted as a Markovian SI epidemic that spreads deterministically locally in space, but also allows for occasional random long range contacts. They showed that the development of the proportion of space infected can be described in an entirely similar way, with an initial branching phase followed by an almost deterministic progression. Barbour & Reinert (2013a) used an approach that they had developed for the analysis of certain small world stochastic networks to extend this result to much more general gossip processes. Their method involves only branching process approximations and asymptotics, and they were able to identify the (otherwise mysterious) expression for the deterministic curve in terms of the Laplace transform of the distribution of the limiting random variable  $W$  for a branching process backwards in time — in this case, having the same distribution as that used for approximating the initial stages of the process. In a subsequent paper, Barbour & Reinert (2013b), they extended their approach to quite general SIR models, with no Markovian assumption; here, the backward and forward branching processes may have different distributions, but the mean measures of the offspring point processes are the same for both. Analogous results for the mean path have recently been independently obtained for the somewhat simpler first passage percolation model by Bhamidi *et. al.* (2012).

Suppose that an epidemic spreads from an initial infective in a population of large size  $N$ . The idea of the method is roughly as follows. A given individual  $K$  has been infected by time  $2t + s$  only if it was infected within time  $t + s$  by some individual that was infected by the initial infective within time  $t$ . The number of individuals that is infected by the initial infective within time  $t = t_N$  can be well approximated using a branching process, if  $t_N$  is not too large, and the same is true of the number of individuals who would infect  $K$  within time  $t_N + s$ , for the same choice of  $t_N$ , though the branching processes may be different. The probability that these two sets of individuals share a common member is then given by a hypergeometric probability, if all individuals are equivalent with respect to contact probability (homogeneous mixing), and, if  $t_N$  is correctly specified, this probability is asymptotically non-trivial. As it happens,  $t_N$  should be chosen in such a way that the mean number of individuals in (either) branching process at time  $t_N$  is about  $\sqrt{N}$ ; if  $\lambda$  is the Malthusian parameter (it is the same for both branching processes), take  $t_N = (1/2\lambda) \log N$ . For this choice, there are about  $c_f W_f \sqrt{N}$  individuals currently infected at time  $t_N$ , and the set of individuals who would infect  $K$  within a further time  $t_N + s$  has size about  $C_b W_b \sqrt{N} e^{\lambda s}$ ; here,  $W_f$  and  $W_b$  are limiting random variables for the forward and backward branching processes, respectively, and  $c_f$  and  $C_b$  are the constants appropriate for the asymptotics of the number currently alive in the forward branching process and the number that have so far been born in the backward branching process, respectively. Conditional on  $W_f$  and  $W_b$ , the hypergeometric probability is then close to  $1 - \exp\{-c_f C_b W_f W_b e^{\lambda s}\}$ . Hence the expectation of the proportion  $\Pi(2t_N + s)$  of

individuals that have *not* been infected by time  $2t_N + s$ , given the initial development up to time  $t_N$ , which is the same as the conditional probability that a randomly chosen  $K$  is not infected, is close to

$$\begin{aligned} \mathbb{E}\left\{e^{-c_f C_b W_f W_b e^{\lambda s}} \mid W_f\right\} &= \psi_B(e^{\lambda s + \log(c_f C_b) + \log W_f}) \\ &= s_B(s + \lambda^{-1}\{\log(c_f C_b) + \log W_f\}), \end{aligned}$$

where  $\psi_B(v) := \mathbb{E}\{e^{-vW_b}\}$  and  $s_B(u) := \psi_B(e^{\lambda u})$ . By a similar argument, the expectation of  $\{\Pi(2t_N + s)\}^2$  is the same as the probability that two independently and uniformly chosen individuals  $K$  and  $K'$  are uninfected at time  $2t_N + s$ , and this in turn is close to  $\{s_B(s + \lambda^{-1}\{\log(c_f C_b) + \log W_f\})\}^2$ , because the numbers of individuals who would infect  $K$  and  $K'$  within time  $t_N + s$  are close to being independent. But this implies that the conditional variance of  $\Pi(2t_N + s)$  is small, and hence that  $\Pi(2t_N + s)$  is asymptotically close to its conditional expectation, given the development up to time  $t_N$ . Note that the determination of the constant  $c_f C_b$  appearing in the approximation actually requires a more careful argument than has been presented in this sketch.

One of the technical difficulties in carrying through this argument precisely is that the branching process coupling, using Ball's (1983) method, may well fail before time  $t_N$ . Then it is necessary to control the effects of 'ghosts', individuals that are present in the branching process, but not in the epidemic. There are not many of them, but showing that their influence is small typically requires a disproportionate effort. Here, we use the fact that, for some epidemic models at least, there is a way of coupling the branching and epidemic processes that has asymptotically small failure probability for times even longer than the choice of  $t_N$  above. This greatly reduces the detailed estimates required in the proof. The couplings that we shall use are those introduced in Barbour & Utev (2004) and in Barbour (2010).

## 2 The Reed–Frost epidemic process

The Reed–Frost epidemic process is a chain binomial model, defined by the following stochastic recursion. For given  $S_0$  and  $I_0$ , the numbers of susceptibles  $S_n$  and infectives  $I_n$  at time  $n$  satisfy

$$\mathcal{L}(I_{n+1} \mid S_n, I_n) = \text{Bi}(S_n, 1 - q^{I_n}); \quad S_{n+1} = S_n - I_{n+1}. \quad (2.1)$$

Here,  $q := 1 - p$ , with  $p$  representing the probability that a contact occurs between a given pair of individuals during a single time unit. If  $I_n = 0$  for some  $n$ , then  $S_{n+j} = S_n$  for all  $j \geq 0$ ; the epidemic has terminated. Note also that  $S_n = S_0 - \sum_{l=1}^n I_l$ , so that the path of the process is determined by the values of  $(I_n, n \geq 0)$ . We shall be interested in situations in which the total population  $N = S_0 + I_0$  is large, and we suppose for definiteness that  $p = p_N = m/(N - 1)$  for some fixed  $m$ , the mean number of contacts that an individual makes in one time unit.

Our main theorem, stated precisely in Theorem 2.3 below, shows that, with high probability when  $N$  is large, the proportion of susceptibles  $N^{-1}S_{R(N)+r}$  stays close to the points  $s(u_N + V_N + r)$  on a fixed continuous, decreasing curve  $s$ , for all  $r \in \mathbb{Z}$ .

Here,  $R(N) \sim \log N / \log m$ , the values  $u_N$  are uniformly bounded, and  $V_N$  is a random variable that depends only on the early evolution of the epidemic process, and converges in distribution as  $N \rightarrow \infty$  to a non-trivial limit  $W$ . The distribution of  $W$  and the function  $s$  can both be deduced from the branching process  $Z = (Z_n, n \geq 0)$  with offspring distribution  $\text{Po}(m)$ . Thus the ‘epidemic curve’ described by the proportion of susceptibles is fixed, but it is traversed in discrete time steps, and with a random time shift.

Our proof of the theorem makes use of the following coupling inequality, that is central to the argument. We show that the Reed–Frost epidemic process can be coupled to  $Z$ , if we take  $Z_0 = I_0$ , in such a way that

$$1 - \mathbb{P}[Z_n = I_n \text{ for } 0 \leq n \leq n(N)] = O(N^{-1} I_0^{3/2} m^{3n(N)/2}) = O\{N^{-1} (\mathbb{E}Z_{n(N)})^{3/2}\}. \quad (2.2)$$

We use (2.2), with  $n(N)$  chosen such that  $m^{n(N)} \asymp N^{7/12}$ , to justify using the branching process for calculations about the epidemic process up to the time when (if ever) about  $N^{7/12}$  individuals have been infected. Because we can do this accurately for so long, we never have to cope with the ‘ghosts’ alluded to in the Introduction.

We begin with a Poisson approximation lemma.

**Lemma 2.1** *For  $M \geq s$  and  $i$  in  $\mathbb{Z}_+$  and for  $0 \leq p = 1 - q \leq 1$ ,*

$$d_{TV}(\text{Bi}(s, 1 - q^i), \text{Po}(Mip)) \leq ip + (Mip)^{-1/2} \{(M - s)ip + si(i - 1)p^2/2\}.$$

*If also  $M - s \geq i$ , then*

$$d_{TV}(\text{Bi}(s, 1 - q^i), \text{Po}(Mip)) \leq \frac{1}{2}(M - s) \sqrt{\frac{ip}{M}} \{3 + 2Mp\}.$$

**Proof:** Since, by Taylor’s theorem,  $(1 - q^i) \leq ip$  and  $0 \leq ip - 1 + q^i \leq \frac{1}{2}i(i - 1)p^2$ , it follows from Barbour, Holst & Janson (1992, Chapter 1, (1.23) and Theorem 1.C(i)) that

$$d_{TV}\{\text{Bi}(s, 1 - q^i), \text{Po}(s(1 - q^i))\} \leq ip,$$

and that

$$d_{TV}\{\text{Po}(s(1 - q^i)), \text{Po}(Mip)\} \leq (Mip)^{-1/2} \{(M - s)ip + si(i - 1)p^2/2\}.$$

For the final simplification, note that the result is trivial for  $i = 0$ , and that, for  $i \geq 1$  and  $M - s \geq i$ ,

$$\sqrt{Mip}/(M - s) \leq \sqrt{Mp}/(M - s) \leq \sqrt{Mp} \leq \frac{1}{2}\{1 + Mp\},$$

and

$$s(i - 1)p/\{2(M - s)\} \leq sp/2 \leq Mp/2,$$

from which the last part follows. ■

We now couple a branching process  $Z$  with offspring distribution  $\text{Po}(Mp)$  to the process of infectious individuals  $I$  in the Reed–Frost epidemic, where  $M \geq S_0$ . Lemma 2.1 is used to compare the stochastic recursion (2.1) to the recursion

$$\mathcal{L}(Z_{n+1} | Z_n) = \text{Po}(MpZ_n), \quad (2.3)$$

for the branching process.

**Lemma 2.2** *If  $M \geq S_0$ , then the Reed–Frost process of (2.1) with  $S_0 + I_0 = N$  and the branching process  $Z$  with offspring distribution  $\text{Po}(Mp)$ , both starting with  $Z_0 = I_0$ , can be coupled so that  $Z_n = I_n$ ,  $0 \leq n \leq n_0$ , with probability of failure at most*

$$c(Mp)I_0^{3/2}M^{-1}(Mp)^{n_0/2}\{(M - S_0) + (Mp)^{n_0}\},$$

where  $x^{-1}c(x)$  is bounded in any interval  $x \geq 1 + \delta$  for  $\delta > 0$ . In particular, if  $m > 1$  and  $\alpha > 0$  are fixed and  $p = m/(N - 1)$ ,  $n_0 \leq \alpha \log N / \log m$  and  $M - S_0 \leq N^\alpha$ , then, as  $N \rightarrow \infty$ , the probability of failure is of order  $O(N^{-1+3\alpha/2})$ , and hence is small for any  $\alpha < 2/3$ , for any fixed  $I_0$ .

**Proof:** Let  $B_l := \sum_{j=0}^l Z_j$  denote the number of individuals ever alive in the branching process up to time  $l$ . Then, if  $I_l = Z_l$  for  $0 \leq l \leq n$ , it follows that  $S_n = N - B_n \geq S_0 - B_n$  and that  $I_n = Z_n$ . Comparing the recursions (2.1) and (2.3), noting that  $M - S_n \geq I_n$  and using Lemma 2.1, it follows that  $Z_{n+1}$  and  $I_{n+1}$  can be coupled in such a way that

$$\mathbb{P}[Z_{n+1} \neq I_{n+1} \mid I_l = Z_l, 0 \leq l \leq n] \leq \frac{1}{2}(M - S_0 + B_n)\sqrt{\frac{Z_n p}{M}}\{3 + 2Mp\}.$$

Write  $\tilde{m} := Mp$ , and note that, for the martingale  $W_n := \tilde{m}^{-n}Z_n$  with limit  $W$ , we have

$$\mathbb{E}\{(W_{n+r} - W_n)^2\} \leq \frac{I_0 m^{-n}}{m-1} \quad \text{and} \quad \mathbb{E}W_n^2 \leq \frac{I_0^2 m}{m-1}. \quad (2.4)$$

It thus follows that  $\mathbb{E}\{\sqrt{Z_n}\} \leq \tilde{m}^{n/2}\sqrt{I_0}$  and that

$$\begin{aligned} \mathbb{E}\{B_n \sqrt{Z_n}\} &= \sum_{r=0}^n m^{r+n/2} \mathbb{E}\{W_r \mathbb{E}(\sqrt{W_n} \mid Z_r)\} \\ &\leq \sum_{r=0}^n m^{r+n/2} \{\mathbb{E}W_r^2\}^{3/4} \leq \left(\frac{m}{m-1}\right)^{7/4} m^{3n/2} I_0^{3/2}. \end{aligned}$$

Hence  $Z_n$  and  $I_n$  can be coupled exactly for  $1 \leq n \leq n_0$  with failure probability at most

$$c(\tilde{m})I_0^{3/2}M^{-1}\tilde{m}^{n_0/2}\{(M - S_0) + \tilde{m}^{n_0}\},$$

for  $c(x) := (\frac{3}{2} + x) \max\{1, x^{3/4}(x-1)^{-7/4}\}$  and with  $\tilde{m} := Mp$ . ■

Note that Lemma 2.2 establishes (2.2).

Now let  $I_0$  be fixed, and take  $S_0 = N - I_0$ ; let  $p = p_N = m/(N - 1)$  as before. For  $0 < \alpha < 2/3$ , define  $n_\alpha(N, I_0)$  such that

$$\theta_\alpha(N)I_0 m^{n_\alpha(N, I_0)} = N^\alpha, \quad \text{where} \quad m^{-1} < \theta_\alpha(N) \leq 1. \quad (2.5)$$

Define  $n_f(N) := n_{7/12}(N, I_0)$  and  $n_b(N) := n_{5/12}(N, 1)$ , and define  $\theta_f(N)$  and  $\theta_b(N)$  correspondingly. Note that  $n_b(N)$  is the time for which we couple the susceptibility process for a single individual to the backward branching process, so that the initial number here is always 1. Then, by Lemma 2.2, the branching process  $Z$  with  $Z_0 = I_0$  and

with offspring distribution  $\text{Po}(m)$  can be coupled exactly to the process  $I$  of infectives for  $n_f(N)$  steps with failure probability at most of order  $O(N^{-1/8})$ . For the main theorem of this section, we condition on the values of  $Z_n$  for  $n \leq n_f(N)$ , denoting this information by  $\mathcal{F}_0$ . We establish the almost deterministic development from time  $n_f(N)$  onwards only on an event in  $\mathcal{F}_0$  of asymptotically high probability. We let  $\mathcal{C}_N$  denote the event that the coupling up to time  $n_f(N)$  is successful, and set  $A_N^{(1)} := \{B_{n_f(N)} \leq N^{2/3}\}$  and  $A_N^{(2)} := \{|\log W_{n_f(N)}| \leq (1/24) \log N\}$ , where, as before,  $W_n := m^{-n} Z_n$  and  $W := \lim_{n \rightarrow \infty} W_n$ . We then set  $A_N^* := \mathcal{C}_N \cap A_N^{(1)} \cap A_N^{(2)}$ . The probability of  $A_N^*$  is close to the probability  $1 - q^{I_0}$  of non-extinction of the branching process  $Z$ , starting with  $I_0$  individuals. The probability of a small epidemic, when the proportion of susceptibles always stays near 1, is close to  $q^{I_0}$ .

**Theorem 2.3** *For the Reed–Frost epidemic process as defined above, and for any  $\varepsilon > 0$ ,*

$$\lim_{N \rightarrow \infty} \mathbb{P} \left[ \sup_{r \in \mathbb{Z}} |N^{-1} S_{n_f(N)+n_b(N)+r} - s(r + \{\log c_N + \log W_{n_f(N)}\} / \log m)| > \varepsilon \mid \mathcal{F}_0 \cap A_N^* \right] = 0,$$

where  $s(u) := \psi(e^{u \log m})$ ,  $\psi(v) := \mathbb{E}\{e^{-vW}\}$  and  $\lim_{N \rightarrow \infty} \mathbb{P}[A_N^*] = \mathbb{P}[W > 0]$ .

**Proof:** As outlined in the introduction, for  $K$  a uniformly distributed element of  $[N] := \{1, 2, \dots, N\}$ ,

$$\begin{aligned} \mathbb{E}\{N^{-1} S_{n_f(N)+n_b(N)+r} \mid \mathcal{F}_0 \cap \mathcal{C}_N\} &= \mathbb{P}[K \in \mathcal{S}_{n_f(N)+n_b(N)+r} \mid \mathcal{F}_0 \cap \mathcal{C}_N] \\ &= \mathbb{P}[K \notin \mathcal{B}_{n_f(N)-1} \mid \mathbb{P}[\tilde{\mathcal{U}}_{n_b(N)+r} \cap \mathcal{Z}_{n_f(N)} = \emptyset \mid \mathcal{F}_0 \cap \mathcal{C}_N \cap \{K \notin \mathcal{B}_{n_f(N)-1}\}]], \end{aligned} \quad (2.6)$$

where  $\mathcal{B}_{n_f(N)-1}$  denotes the set of indices randomly assigned to the  $B_{n_f(N)-1}$  individuals infected before time  $n_f(N)$ ,  $\mathcal{Z}_{n_f(N)}$  those assigned to the  $Z_{n_f(N)}$  infected individuals at time  $n_f(N)$ ,  $\mathcal{S}_l$  those assigned to individuals uninfected at time  $l$  and  $\tilde{\mathcal{U}}_{n_b(N)+r}$  those assigned to individuals who would contact  $K$  within an elapsed time of  $n_b(N) + r$ . If  $K \in \mathcal{B}_{n_f(N)-1}$ ,  $K$  cannot belong to  $\mathcal{S}_{n_f(N)+n_b(N)+r}$ , since it has already been infected before time  $n_f(N)$ . If not, then no member of  $\tilde{\mathcal{U}}_{n_b(N)+r}$  can belong to  $\mathcal{B}_{n_f(N)-1}$  unless there is also a member of  $\tilde{\mathcal{U}}_{n_b(N)+r}$  that belongs to  $\mathcal{Z}_{n_f(N)}$ , because all individuals directly in contact with  $\mathcal{B}_{n_f(N)-1}$  belong to  $\mathcal{B}_{n_f(N)-1} \cup \mathcal{Z}_{n_f(N)}$ , by construction. Thus, for  $\tilde{\mathcal{U}}_{n_b(N)+r} := |\tilde{\mathcal{U}}_{n_b(N)+r}|$ , we have

$$\begin{aligned} \mathbb{P}[\tilde{\mathcal{U}}_{n_b(N)+r} \cap \mathcal{Z}_{n_f(N)} = \emptyset \mid \mathcal{F}_0 \cap \mathcal{C}_N \cap \{K \notin \mathcal{B}_{n_f(N)-1}\}; \tilde{\mathcal{U}}_{n_b(N)+r}] \\ = P(N - B_{n_f(N)-1}, Z_{n_f(N)}, \tilde{\mathcal{U}}_{n_b(N)+r}), \end{aligned} \quad (2.7)$$

where  $P(n, m_1, m_2)$  denotes the hypergeometric probability that two independently chosen uniform random subsets of  $[n]$  of sizes  $m_1$  and  $m_2$  do not intersect.

Now the  $\tilde{U}_l$ ,  $l \geq 0$ , again arise as a Reed–Frost epidemic, this time on the set  $[N] \setminus \mathcal{B}_{n_f(N)-1}$  and with the same  $p$  as before; hence we can use Lemma 2.2 once more to show that they can be generated by a branching process  $(\tilde{Z}_l, l \geq 0)$  with  $\tilde{Z}_0 = 1$  and with offspring distribution  $\text{Po}(m)$ , with exact coupling failing to hold up to time  $n_b(N) + r$  with probability of order  $O(N^{-1/12})$ , on the set  $A_N^{(1)}$ , if  $r$  is such that  $m^r \leq N^{1/12}$ ; we write  $\tilde{B}_l$

to denote  $\sum_{j=0}^l \tilde{Z}_j$  and  $\tilde{W} := \lim_{n \rightarrow \infty} m^{-n} \tilde{Z}_n$ . Noting also that  $\mathbb{P}[K \in \mathcal{B}_{n_f(N)-1}] \leq N^{-1/3}$  on  $A_N^{(1)}$ , it follows from (2.6), and by taking expectation with respect to  $\tilde{U}_{n_b(N)+r}$  in (2.7), that

$$\begin{aligned} & \mathbb{E}\{N^{-1}S_{n_f(N)+n_b(N)+r} \mid \mathcal{F}_0 \cap \mathcal{C}_N \cap A_N^{(1)}\} \\ &= \mathbb{E}\{P(N - B_{n_f(N)-1}, Z_{n_f(N)}, \tilde{B}_{n_b(N)+r}) \mid \mathcal{F}_0 \cap \mathcal{C}_N \cap A_N^{(1)}\} + O(N^{-1/12}). \end{aligned} \quad (2.8)$$

Furthermore,

$$0 \leq \exp\{-m_1 m_2 / n\} - P(n, m_1, m_2) \leq n^{-1}(m_1 + m_2), \quad (2.9)$$

by Barbour, Holst & Janson (1992, Theorem 6A), and, for  $0 \leq x \leq m$  and  $0 \leq y \leq n/2$ ,

$$0 \leq e^{-(m-x)l/n} - e^{-ml/n} \leq \frac{xl}{n}; \quad 0 \leq e^{-ml/n} - e^{-ml/(n-y)} \leq 4 \frac{mly}{n^2} e^{-ml/n}. \quad (2.10)$$

Hence

$$\begin{aligned} & |\mathbb{E}\{N^{-1}S_{n_f(N)+n_b(N)+r} \mid \mathcal{F}_0 \cap \mathcal{C}_N \cap A_N^{(1)}\} - \mathbb{E}\{\exp(-N^{-1}Z_{n_f(N)}\tilde{B}_{n_b(N)+r}) \mid \mathcal{F}_0 \cap \mathcal{C}_N\}| \\ & \leq C'(N^{-1/3} + N^{-7/12} + N^{-1/12}), \end{aligned} \quad (2.11)$$

for a constant  $C'$ .

Now, from (2.4) with  $I_0$  replaced by 1, if  $r$  is such that  $m^r \geq N^{-1/12}$ ,

$$\mathbb{E}|m^{-(n_b(N)+r)}\tilde{Z}_{n_b(N)+r} - \tilde{W}| = O(N^{-1/6}).$$

Similarly, also using (2.4) with  $I_0 = 1$ ,

$$\begin{aligned} \mathbb{E}\{(\tilde{B}_{n-1} - \tilde{Z}_n/(m-1))^2\} &= \mathbb{E}\left\{\left(\sum_{r=0}^{n-1} m^r (\tilde{W}_r - \tilde{W}_n) - \frac{\tilde{W}_n}{m-1}\right)^2\right\} \\ &\leq (n+1) \left\{ \sum_{r=0}^{n-1} m^{2r} \mathbb{E}\{(\tilde{W}_r - \tilde{W}_n)^2\} + \frac{m}{(m-1)^3} \right\} \\ &\leq 2m^n \frac{m(n+1)}{(m-1)^3}, \end{aligned}$$

so that, in the same range of  $r$ ,

$$\mathbb{E}|m^{-(n_b(N)+r)}\{\tilde{B}_{n_b(N)+r} - m\tilde{Z}_{n_b(N)+r}/(m-1)\}| = O(N^{-1/6} \log N).$$

Thus we conclude that, for  $|r \log m| \leq (1/12) \log N$ ,

$$\begin{aligned} & |\mathbb{E}\{\exp(-N^{-1}Z_{n_f(N)}\tilde{B}_{n_b(N)+r}) \mid \mathcal{F}_0 \cap \mathcal{C}_N\} \\ & - \mathbb{E}\{-(m-1)I_0\theta_f(N)\theta_b(N))^{-1}m^{r+1}W_{n_f(N)}\tilde{W} \mid \mathcal{F}_0 \cap \mathcal{C}_N\}| = O(N^{-1/12}), \end{aligned} \quad (2.12)$$

since also, from (2.5),  $N = I_0\theta_f(N)\theta_b(N)m^{n_f(N)+n_b(N)}$ . Combining (2.11) and (2.12), it follows that

$$\mathbb{E}\{N^{-1}S_{n_f(N)+n_b(N)+r} \mid \mathcal{F}_0 \cap \mathcal{C}_N \cap A_N^{(1)}\} = \psi(c_N W_{n_f(N)} m^r) + O(N^{-1/12}), \quad (2.13)$$

where  $c_N := m/\{(m-1)I_0\theta_f(N)\theta_b(N)\}$ , uniformly in  $|r \log m| \leq (1/12) \log N$ , and  $W_{n_f(N)}$  can be replaced by  $W$  using (2.4) once more, without change to the order of the error.

It remains to examine  $\text{Var} \{N^{-1}S_{n_f(N)+n_b(N)+r} | \mathcal{F}_0 \cap \mathcal{C}_N \cap A_N^{(1)}\}$ . As above, we can write

$$\mathbb{E}\{(N^{-1}S_{n_f(N)+n_b(N)+r})^2 | \mathcal{F}_0 \cap \mathcal{C}_N \cap A_N^{(1)}\} = \mathbb{P}[\{K, K'\} \cap \mathcal{S}_{n_f(N)+n_b(N)+r} = \emptyset | \mathcal{F}_0 \cap \mathcal{C}_N \cap A_N^{(1)}], \quad (2.14)$$

for two points  $K^{(1)}$  and  $K^{(2)}$  chosen independently and uniformly from  $[N]$ . Once again, only the event  $E_N^{(1)} := \{\{K^{(1)}, K^{(2)}\} \cap \mathcal{B}_{n_f(N)-1} = \emptyset\}$  is important, since the complementary probability is of order  $O(N^{-1/3})$  on  $A_N^{(1)}$ . Then, the corresponding sets  $\tilde{\mathcal{U}}_{n_b(N)+r}^{(i)}$ ,  $i = 1, 2$ , can be constructed by using independent branching processes  $\tilde{Z}^{(i)}$ , with an error of order  $O(N^{-1/12})$ . Note that the probability of the event  $(E_N^{(2)})^c$  that  $\tilde{\mathcal{U}}_{n_b(N)+r}^{(1)}$  and  $\tilde{\mathcal{U}}_{n_b(N)+r}^{(2)}$  intersect, conditional on their sizes  $\tilde{U}_{n_b(N)+r}^{(i)}$  being known, is the hypergeometric probability

$$1 - P(N - B_{n_f(N)-1}, \tilde{U}_{n_b(N)+r}^{(1)}, \tilde{U}_{n_b(N)+r}^{(2)}) = O(N^{-1}\tilde{U}_{n_b(N)+r}^{(1)}\tilde{U}_{n_b(N)+r}^{(2)})$$

on  $A_N^{(1)}$ . Taking expectations, it follows that

$$\mathbb{P}[(E_N^{(2)})^c | \mathcal{F}_0 \cap \mathcal{C}_N \cap A_N^{(1)} \cap E_N^{(1)}] = O(N^{-1/12}).$$

But then, given the sizes  $\tilde{U}_{n_b(N)+r}^{(i)}$  and that the two sets do not intersect, we have

$$\begin{aligned} & \mathbb{P}[\{K, K'\} \cap \mathcal{S}_{n_f(N)+n_b(N)+r} = \emptyset | \mathcal{F}_0 \cap \mathcal{C}_N \cap A_N^{(1)} \cap E_N^{(1)} \cap E_N^{(2)}; \tilde{U}_{n_b(N)+r}^{(i)}, i = 1, 2] \\ &= P(N - B_{n_f(N)-1}, Z_{n_f(N)}, \tilde{U}_{n_b(N)+r}^{(1)})P(N - B_{n_f(N)-1} - \tilde{U}_{n_b(N)+r}^{(1)}, Z_{n_f(N)}, \tilde{U}_{n_b(N)+r}^{(2)}), \end{aligned}$$

and, on  $A_N^{(1)}$  and if also  $\tilde{U}_{n_b(N)+r}^{(1)} \leq N^{2/3}$ , this is equal to

$$\begin{aligned} & P(N - B_{n_f(N)-1}, Z_{n_f(N)}, \tilde{U}_{n_b(N)+r}^{(1)})P(N - B_{n_f(N)-1}, Z_{n_f(N)}, \tilde{U}_{n_b(N)+r}^{(2)}) \\ &+ O(N^{-2}(\tilde{U}_{n_b(N)+r}^{(1)}\tilde{U}_{n_b(N)+r}^{(2)}Z_{n_f(N)})), \end{aligned}$$

since  $|P(N, r, s) - P(N - K, r, s)| \leq rsK/\{N(N - K)\}$ . Replacing  $\tilde{U}_{n_b(N)+r}^{(i)}$  by  $\tilde{B}_{n_b(N)+r}^{(i)}$  and taking expectations, and noting the independence of  $\tilde{B}^{(1)}$  and  $\tilde{B}^{(2)}$ , this gives

$$\begin{aligned} & \mathbb{P}[\{K, K'\} \cap \mathcal{S}_{n_f(N)+n_b(N)+r} = \emptyset | \mathcal{F}_0 \cap \mathcal{C}_N \cap A_N^{(1)} \cap E_N^{(1)}] \\ &= (\mathbb{E}\{P(N - B_{n_f(N)-1}, Z_{n_f(N)}, \tilde{B}_{n_b(N)+r}^{(1)}) | \mathcal{F}_0 \cap \mathcal{C}_N\})^2 + O(N^{-1/12}), \end{aligned}$$

which in turn implies the same estimate for the probability without conditioning on  $E_N^{(1)}$ . This, in view of (2.14) and (2.8), shows that  $\text{Var} \{N^{-1}S_{n_f(N)+n_b(N)+r} | \mathcal{F}_0 \cap \mathcal{C}_N \cap A_N^{(1)}\}$  is of order  $O(N^{-1/12})$ , for any fixed  $r$ .

The extension to the supremum over  $r$  is almost immediate, because both functions in the approximation are bounded and decreasing. The function  $s(u)$  decreases smoothly



from 1 to the extinction probability  $q$  of the branching process  $\tilde{Z}$ , so that, for any  $\varepsilon > 0$ , there are values  $a_\varepsilon < b_\varepsilon$  such that  $s(a_\varepsilon) = 1 - \varepsilon/2$  and  $s(b_\varepsilon) = \varepsilon/2$ . Hence, given the value of  $W_{n_f(N)}$ , there are no more than  $(b_\varepsilon - a_\varepsilon) + 3$  values of  $r$  such that

$$r + \{\log(c_N) + \log W_{n_f(N)}\} / \log m \in (a_\varepsilon - 1, b_\varepsilon + 1).$$

We have proved  $\varepsilon/2$ -approximation with high probability at each of these points, for  $N$  large enough, provided always that, for each of them,  $|r| \log m \leq (1/12) \log N$ , and this is also the case, for  $N$  large enough, if  $|\log W_{n_f(N)}| \leq (1/24) \log N$ . The  $\varepsilon$ -approximation for values of  $r$  outside this range then follows by monotonicity.  $\blacksquare$

The statement of the theorem is complicated by the fact that the process is significantly discrete in the range in which important change in the proportion of susceptibles takes place. The underlying function  $s$  is the same for each  $N$ , but it is observed only at arguments lying on a lattice of span 1. The position of this lattice on the line is random, depending on the value taken by the limiting branching random variable  $W$  relevant to the early development of the process. There is also a deterministic offset  $\log c_N / \log m$ , which varies with  $N$ , and is again a feature of the discrete generation structure in the process.

To interpret the function  $s$ , note that writing  $s_n := N^{-1}S_n$  in (2.1) gives the large  $N$  approximation

$$s_n - s_{n+1} \approx s_n \left( 1 - \left( 1 - \frac{m}{N} \right)^{N(s_{n-1} - s_n)} \right) \approx s_n (1 - e^{-m(s_{n-1} - s_n)}). \quad (2.15)$$

On the other hand, the Laplace transform  $\psi$  of the branching process limiting random variable  $W$  satisfies  $\psi(v) = f(\psi(v/m))$  with  $f(z) = e^{-m(1-z)}$ , so that, for any  $c \in \mathbf{R}$ ,

$$s(n+1+c) = \psi(m^{n+1+c}) = e^{-m(1-\psi(m^{n+1+c}))} = e^{-m(1-s(n+c))},$$

and this gives

$$s(n+c) - s(n+1+c) = s(n+c)(1 - e^{-m(s(n-1+c) - s(n+c))}),$$

which, with  $s(r+c)$  replaced by  $s_r$ , recovers (2.15) with equality.

### 3 The Markovian SIR epidemic process

The second example is the Markovian SIR epidemic process, formulated by Bartlett (1949). The process is a continuous time pure jump Markov process  $\{(S_N(t), I_N(t)), t \geq 0\}$  on  $(\mathbb{Z}_+)^2$ , starting with  $(S_N(0), I_N(0)) = (N, I_0)$ , and having transition rates

$$(S, I) \rightarrow (S-1, I+1) \text{ at rate } \alpha I(S/N); \quad (S, I) \rightarrow (S, I-1) \text{ at rate } \mu I, \quad (S, I) \in (\mathbb{Z}_+)^2,$$

where  $\alpha$  is the *per capita* contact rate of an infective, and  $\mu$  is the recovery rate. We shall think of  $N$  as being large, and take  $I_0$  to be fixed; we shall also assume that  $\alpha > \mu$ , so that the chance of a significant epidemic is not small. Then it has long been understood

that  $(S_N, I_N)$  can be approximated by an initial birth and death phase, the Whittle (1955) approximation, until the random time  $\tau_N := \tau_N^\gamma$  at which the number of infectives has reached some chosen power  $N^\gamma$  of  $N$ , for some suitable  $0 < \gamma < 1$ . After this, the development of  $N^{-1}(S_N(t), I_N(t))$  is well described by the solution of the differential equations

$$\frac{ds}{dt} = -\alpha s(t)i(t); \quad \frac{di}{dt} = (\alpha s(t) - \mu)i(t) \quad (3.1)$$

in  $t \geq \tau_N$ , with  $s(\tau_N) = N^{-1}S_N(\tau_N)$  and  $i(\tau_N) = N^{-1+\gamma}$ . We now illustrate how our approach can be used to justify these approximations, with  $\gamma = 7/12$ ; the formal statement is given in Theorem 3.4. As in the previous section, a key ingredient in the proof is the accurate coupling of the epidemic and branching processes until the first time (if ever) that there are  $N^\gamma$  infectives, where  $1/2 < \gamma < 2/3$ . A more detailed description of the final phase, when  $I_N(t)$  falls below  $N^\gamma$  again, can also be undertaken if, for instance, the distribution of the time until the epidemic finally dies out is of interest: see Barbour (1975).

The initial branching approximation is made using the process  $(S_N^*, I^*)$  having  $N$ -independent transition rates

$$(S, I) \rightarrow (S - 1, I + 1) \text{ at rate } \alpha I; \quad (S, I) \rightarrow (S, I - 1) \text{ at rate } \mu I, \quad (S, I) \in \mathbb{Z} \times \mathbb{Z}_+,$$

with  $(S_N^*(0), I^*(0)) = (N, I_0)$ . Note that the distribution of the second component  $I^*$  is indeed the same for all  $N$ , as is that of  $B^* := N - S_N^*$ , and that the paths of  $B^*$  can be deduced from those of  $I^*$ .  $I^*$  itself is a linear birth and death process with constant *per capita* birth and death rates  $\alpha$  and  $\mu$ .

**Theorem 3.1** *Let  $\tau_N^\gamma := \inf\{t \geq 0: I_N(t) \geq N^\gamma\}$ , with  $\tau_N^\gamma = \infty$  if there is no such  $t$ . Then the epidemic process  $(S_N, I_N)$  can be coupled to a process  $(S_N^*, I^*)$  as defined above, in such a way that*

$$p_N^C(\gamma) := \mathbb{P}[(S_N(t), I_N(t)) = (S_N^*(t), I^*(t)) \text{ for all } 0 \leq t \leq \tau_N^\gamma] = 1 - O(N^{-1+3\gamma/2}).$$

*In particular,  $\lim_{N \rightarrow \infty} p_N^C(\gamma) = 1$  if  $\gamma < 2/3$ .*

**Proof:** The proof can be carried out along exactly the same lines as the Radon–Nikodym argument in Theorem 3.1 of Barbour (2010), though much more simply, to arrive at the statement

$$\mathbb{P}[(S_N(t), I_N(t)) = (S_N^*(t), I^*(t)) \text{ for all } 0 \leq t \leq T_m] \geq 1 - CN^{-1}m^{3/2},$$

for any  $m \in \mathbb{Z}_+$ , where  $T_m$  denotes the time of the  $m$ -th jump, and  $C$  is a constant. However, because the jump chain of  $I^*$  is a simple random walk, its position after  $m$  jumps can be written as  $I_0 + 2Z_m - m$ , where  $Z_m \sim \text{Bi}(m, p)$  and  $p := \alpha/(\alpha + \mu) > 1/2$ . Hence, taking  $m = m(N) = \lceil 2N^\gamma/(2p - 1) \rceil$ , we have  $\mathbb{P}[I_0 + 2Z_m - m < N^\gamma] = o(N^{-r})$  for any  $r > 0$ , by the Chernoff inequality for the binomial distribution, which entails  $\mathbb{P}[\tau_N^\gamma > T_{m(N)}] = o(N^{-r})$  also. This completes the proof. ■

The backward (susceptibility) process is more delicate. Suppose that the set of individuals born up to time  $\tau_N^\gamma$  in the epidemic is denoted by  $\mathcal{B}_N$ , for the fixed value

$\gamma = \gamma_0 := 7/12$ , and write  $B_N := |\mathcal{B}_N|$ , which is also equal to  $B^*(\tau_N^{\gamma_0})$ , if the coupling is successful. Consider the susceptibility process for an individual  $K_0 \notin \mathcal{B}_N$ . Then the probability that it would be infected by any other individual  $K' \notin \mathcal{B}_N$ , if that individual were ever infected, is  $\alpha/(N\mu + \alpha)$ . This is the probability that a contact would take place before the end of  $K'$ 's infectious period, when the contact rate is  $\alpha/N$  and the infectious period has the  $\text{Exp}(\mu)$  distribution. Conditional on this unlikely event, the distribution of the time to infection is  $\text{Exp}(\mu + \alpha/N)$ . Thus the ‘offspring’ point process of  $K_0$  in its susceptibility process is a sum of  $N - B_N - 1$  independent point processes, each of which has no point with probability  $N\mu/(N\mu + \alpha)$ , and a single point with position having the  $\text{Exp}(\mu + \alpha/N)$  distribution on the complementary event. By the Poisson approximation to the binomial, this process differs from a Poisson process with intensity  $\rho_N \alpha e^{-(\mu + \alpha/N)u}$  at  $u \in \mathbf{R}_+$  only on an event of probability of order  $O(N^{-1})$ ; here,  $\rho_N := 1 - (B_N + 1)/N$ .

Now consider the offspring of the  $(n + 1)$ 'st individual  $K_n$  of the susceptibility process, in order of ‘birth’ time. There are  $n$  individuals, including  $K_0$ , that have already been considered, and they have had  $U_n$  offspring, say. The probability of any one of these contacting  $K_n$  is of order  $(n + U_n)/N$ . The remaining individuals in  $[N] \setminus \mathcal{B}_N$  are such that their  $\text{Exp}(\mu)$ -distributed infectious periods were shorter than the independent and  $\text{Exp}(\alpha/N)$ -distributed times to the first points in the infection processes along the links from them to individuals  $K_0, K_1, \dots, K_{n-1}$ . As a result, the probability of such a  $K'$  infecting  $K_n$  is now  $\alpha/(N\mu + (n + 1)\alpha)$ , and, should it do so, the distribution of the time until infection is  $\text{Exp}(\mu + (n + 1)\alpha/N)$ . Hence, except on an event of probability of order  $O(N^{-1}(n + U_n))$ , the offspring point process for  $K_n$  has the distribution of a Poisson process with intensity  $\rho_N \alpha e^{-(\mu + (n+1)\alpha/N)u}$  on  $\mathbf{R}_+$ . We now show that a Poisson process with this intensity can be replaced by one with intensity  $\rho_N \alpha e^{-\mu u}$ , with only a small error probability, and then that the branching process with this distribution for the offspring point processes differs from that with intensity  $\alpha e^{-\mu u} du$  only on an event of small probability, over a suitable length of time. Since we only need to follow the processes until about  $N^{5/12}$  births have occurred, the cumulative probabilities can be shown to be small.

**Lemma 3.2** *For fixed  $\alpha, \mu > 0$ , let  $\Pi$  and  $\Xi$  be Poisson processes with intensities  $\alpha e^{-\mu u}$  and  $\alpha e^{-\nu u}$ , respectively. Then, uniformly in  $\nu \geq \mu$ ,*

$$d_{TV}(P, Q) = O(|1 - \mu/\nu|),$$

where  $P := \mathcal{L}(\Pi)$ ,  $Q := \mathcal{L}(\Xi)$ .

**Proof:** The processes  $\Pi$  and  $\Xi$  each have a.s. only finitely many points. The likelihood of a realization of  $\Pi$  with  $k$  points at times  $t_1 < t_2 < \dots < t_k < \infty$  is

$$\exp(-\alpha e^{-\mu t_k}/\mu) \prod_{j=1}^k \{\alpha e^{-\mu t_j} \exp(-\alpha(e^{-\mu t_{j-1}} - e^{-\mu t_j})/\mu)\} = e^{-\alpha/\mu} \alpha^k e^{-\mu s_k},$$

where  $s_k := t_1 + \dots + t_k$  and  $t_0 := 0$  in the product. Hence

$$\frac{dQ}{dP}(t_1, \dots, t_k) = e^{-(\mu - \nu)\{\alpha/(\mu\nu) - s_k\}}.$$

Now, letting  $T_j \leq \infty$  denote the time of the  $j$ -th point of a random realization under  $\Pi$ , and  $S := \sum_{j:T_j < \infty} T_j$ ,

$$\mathbb{E}_P\{e^{\theta S}\} = \exp\left\{\alpha \int_0^\infty e^{-\mu u}(e^{\theta u} - 1) du\right\} = \exp\left\{\frac{\alpha\theta}{\mu(\mu - \theta)}\right\},$$

for any  $\theta < \mu$ . This implies, taking  $\theta = -2(\nu - \mu)$ , that

$$\begin{aligned} \mathbb{E}_P\left\{\left(\frac{dQ}{dP}((T_j, j \geq 1))\right)^2\right\} &= \exp\{2\alpha(\nu - \mu)/(\mu\nu)\}\mathbb{E}_P\{e^{-2(\nu - \mu)S}\} \\ &= \exp\{2\alpha(\nu - \mu)^2/(\mu\nu(2\nu - \mu))\}. \end{aligned}$$

It is then immediate that  $\text{Var}_P\{dQ/dP\} = O((1 - \mu/\nu)^2)$ , uniformly in  $\nu \geq \mu$ , and the lemma follows because, if  $P$  and  $Q$  are two probability measures on the same space, and  $X \sim P$ , then

$$2d_{TV}(P, Q) = \mathbb{E}\left|\frac{dQ}{dP}(X) - 1\right|.$$

■

**Lemma 3.3** *Let  $Z_\alpha$  denote a Crump–Mode–Jagers branching process with  $Z_\alpha(0) = 1$ , whose offspring point processes are distributed as Poisson processes  $P_a$  with intensities  $ae^{-\mu u}$  at  $u \in \mathbf{R}_+$ . Then, for  $\alpha_1 < \alpha_2 < 3\alpha_1/2$  and for any  $m \geq 1$ ,*

$$d_{TV}\{\mathcal{L}(Z_{\alpha_1}(s), 0 \leq s \leq \tau_{m+1}^{(1)}), \mathcal{L}(Z_{\alpha_2}(s), 0 \leq s \leq \tau_{m+1}^{(2)})\} \leq \sqrt{9em/8}|1 - \alpha_1/\alpha_2|,$$

where  $\tau_m^{(l)}$  denotes the birth time of the  $m$ -th individual in  $Z_{\alpha_l}$ .

**Proof:** For  $x > 0$  and  $f$  an integrable non-negative function on  $\mathbf{R}_+$  with  $F(t) := \int_0^t f(u) du$ , consider the probability measure  $Q_x$  with density  $xf(u)e^{-xF(u)}$  at  $u \in \mathbf{R}_+$ , and with mass  $e^{-xF(\infty)}$  on  $+\infty$ . Then, for  $a, b > 0$ ,

$$1 \leq \mathbb{E}_{Q_b}\left\{\left(\frac{dQ_a}{dQ_b}(U)\right)^2\right\} = 1 + \frac{(a-b)^2}{b(2a-b)}\{1 - e^{-(2a-b)F(\infty)}\} \leq 1 + \frac{(a-b)^2}{b(2a-b)}.$$

Now the likelihood  $dZ_{\alpha_2}/dZ_{\alpha_1}$  evaluated at a path with its first  $k$  jumps at times  $t_1 \leq t_2 \leq \dots \leq t_k \leq \infty$  takes the form of a product  $W_k := \prod_{j=0}^{k-1} V_j$ , where  $V_j$  is a ratio  $(dQ_{\alpha_2}/dQ_{\alpha_1})(t_{j+1})$  as above, for which

$$f_j(u) := A_j e^{-\mu(u-t_j)}, \quad u > t_j, \quad \text{and} \quad A_j := \frac{1}{\mu} \sum_{l=0}^j e^{-\mu(t_j-t_l)}.$$

Hence  $\mathbb{E}\{W_{j+1} | \mathcal{F}_j\} = W_j$  and

$$\mathbb{E}_{\alpha_1}\{(W_{j+1} - W_j)^2 | \mathcal{F}_j\} \leq \{(W_j^2 - 1) + 1\}\varepsilon,$$

where

$$\varepsilon := \frac{(\alpha_2 - \alpha_1)^2}{\alpha_1(2\alpha_2 - \alpha_1)} \leq \frac{9}{8} \left(1 - \frac{\alpha_1}{\alpha_2}\right)^2,$$

for  $\alpha_1 < \alpha_2 \leq 3\alpha_1/2$ . From this it follows, for any  $j \geq 1$ , that

$$\text{Var}_{\alpha_1} W_{j+1} \leq \text{Var}_{\alpha_1} W_j(1 + \varepsilon) + \varepsilon \leq j\varepsilon e^{j\varepsilon}, \quad (3.2)$$

and hence  $\text{Var}_{\alpha_1} W_m \leq m\varepsilon e$  when  $m\varepsilon \leq 1$ . The lemma is now immediate, since the bound exceeds 1 if  $m\varepsilon > 1$ .  $\blacksquare$

As a result of Lemmas 3.2 and 3.3, and because  $\mathbb{E}U_r \leq r\alpha/\mu$ , it follows that the susceptibility process from any  $K'$ , chosen uniformly at random from  $[N] \setminus \mathcal{B}_N$ , can be replaced by a branching process  $\tilde{Z}$  with the distribution of  $Z_\alpha$  as in Lemma 3.3, until the birth of the  $m + 1$ -st individual, except on an event of probability

$$O\left(\sqrt{m}N^{-1}(B_N + 1) + \sum_{r=0}^m N^{-1}(r + 1 + \mathbb{E}U_r)\right) = O(\sqrt{m}N^{-1}(B_N + 1) + m^2N^{-1}),$$

and, if  $B_N \leq N^{2/3}$ , this is asymptotically small whenever  $m = m_N = o(N^{1/2})$ . Now, since  $Z_\alpha$  has Malthusian parameter  $\lambda := \alpha - \mu > 0$  and  $\mathbb{E}\left\{\left(\int_0^\infty e^{-\lambda u} P_\alpha(du)\right)^2\right\} < \infty$ , it follows that  $C := \sup_t \mathbb{E}\{Z_\alpha(t)e^{-(\alpha-\mu)t}\} < \infty$ . Hence, defining

$$t_N(u) := \frac{1}{(\alpha - \mu)} \frac{5}{12} \log N + u, \quad (3.3)$$

and defining  $m_N := N^{23/48}$  and  $u_N := \{24(\alpha - \mu)\}^{-1} \log N$ , it follows by Markov's inequality that

$$\mathbb{P}[Z_\alpha(t_N(u_N)) > m_N] \leq CN^{-1/48}.$$

Thus, on  $\{B_N \leq N^{2/3}\}$ , the susceptibility process from  $K'$  can be coupled to  $Z_\alpha$  up to time  $t_N(u_N)$  with failure probability of order  $O(N^{-1/48})$ . This enables us to prove the following theorem. The theorem cannot be deduced from the results of Barbour & Reinert (2013b), because the exponential distribution of an individual's infectious period does not satisfy their Assumption 2: see their Remark 2.1.

**Theorem 3.4** *For the Markovian SIR epidemic process as defined above, and for any  $\varepsilon > 0$ ,*

$$\lim_{N \rightarrow \infty} \mathbb{P}\left[\sup_{u \in \mathbf{R}} |N^{-1}S(\tau_N^{\gamma_0} + t_N(u)) - s(u)| > \varepsilon \mid \mathcal{F}_0 \cap \{\tau_N^{\gamma_0} < \infty\} \cap \{B_N \leq N^{2/3}\}\right] = 0,$$

where  $\gamma_0 := 7/12$ ,  $\mathcal{F}_0 := \mathcal{F}_{\tau_N^{\gamma_0}}$ ,  $s(u) := \psi(e^{(\alpha-\mu)u})$  and  $\psi(v) := \mathbb{E}\{e^{-v\tilde{W}}\}$ . Note that  $\lim_{N \rightarrow \infty} \mathbb{P}[\{\tau_N^{\gamma_0} < \infty\} \cap \{B_N > N^{2/3}\}] = 0$  and  $\lim_{N \rightarrow \infty} \mathbb{P}[\tau_N^{\gamma_0} < \infty] = \mathbb{P}[W^* > 0]$ . Here,  $W^* := \lim_{t \rightarrow \infty} W^*(t)$  and  $W^*(t) := e^{-(\alpha-\mu)t} I^*(t)$ ;  $\tilde{W}$  denotes the limiting random variable  $\lim_{t \rightarrow \infty} e^{-(\alpha-\mu)t} \tilde{Z}(t)$  for the backward branching process.

**Proof:** The proof is similar in spirit to that of Theorem 2.3. At time  $t = \tau_N^{\gamma_0}$ ,  $I_N(t) = I^*(t) = N_{\gamma_0} := \lceil N^{\gamma_0} \rceil$  and  $S_N(t) = S_N^*(t)$  with probability at least  $1 - O(N^{-1/8})$ , from Theorem 3.1. From the argument preceding the statement of the theorem, the backward process  $\tilde{Z}$  can be taken to be distributed as  $Z_\alpha$  up to time  $t_N(u)$ , with error probability of order  $O(N^{-1/48})$ , for any  $u \leq u_N$ , and at time  $t = t_N(u)$  it has  $\tilde{Z}_t = \tilde{W}_t N^{5/12} e^{(\alpha-\mu)u}$  individuals.

Let the  $j$ -th of the  $N_{\gamma_0}$  individuals  $\mathcal{B}_{\gamma_0}$  of  $I_N$  at time  $\tau_N^{\gamma_0}$  have  $C_j$  direct offspring born at times  $(\tau_{jl}, 1 \leq l \leq C_j)$ . Then a random point  $K \in [N]$  is such that  $K \notin \mathcal{S}(\tau_N^{\gamma_0} + t_N(u))$  if  $K \notin \mathcal{B}_{\gamma_0}$ , and if

$$V_N(u) := \sum_{j=1}^{N_{\gamma_0}} \sum_{l=1}^{C_j} \sum_{r=1}^{\tilde{Z}_{t_N(u)}} I[L_{jl} = \tilde{L}_r] I[\tau_{jl} \leq A_r] = 0,$$

where  $(L_{jl}, 1 \leq l \leq C_j, 1 \leq j \leq N_{\gamma_0})$  denote the distinct labels assigned to the direct offspring of the individuals in  $\mathcal{B}_{\gamma_0}$ , and  $(\tilde{L}_r, 1 \leq r \leq \tilde{Z}_{t_N(u)})$  those corresponding to the individuals in the version of  $\tilde{Z}$  corresponding to  $K$ ;  $A_r$  is the age of the  $r$ -th individual of  $\tilde{Z}$  at time  $t_N(u)$ . The probability of the latter event is approximated by  $e^{-\mathbb{E}V_N(u)}$  with error at most of order  $O(N^{-5/12})$  by Barbour, Holst & Janson (1992, Theorem 4.A), where

$$\mathbb{E}V_N(u) = \tilde{W}_{t_N(u)} e^{(\alpha-\mu)u} + O(N^{-5/24}),$$

because

$$\mathbb{P}[A_r \geq v] = \tilde{W}_{t_N(u)-v} e^{-(\alpha-\mu)v} / \tilde{W}_{t_N(u)} = e^{-(\alpha-\mu)v} \{1 + O(N^{-5/24})\},$$

and

$$\mathbb{E} \left\{ \sum_{l=1}^{C_j} e^{-(\alpha-\mu)\tau_{jl}} \right\} = 1.$$

Thus  $s(u)$  approximates  $N^{-1} \mathbb{E}\{S(\tau_N^{\gamma_0} + t_N(u)) \mid \mathcal{F}_0 \cap \{\tau_N^{\gamma_0} < \infty\} \cap \{B_N \leq N^{2/3}\}\}$ , and the rest of the proof is then much as for Theorem 2.3. That  $\lim_{N \rightarrow \infty} \mathbb{P}[\{\tau_N^{\gamma_0} < \infty\} \cap \{B_N > N^{2/3}\}] = 0$  follows because, on  $W^* > 0$ ,  $B^*(t)/I^*(t) \rightarrow \alpha/(\alpha-\mu)$  a.s., and hence  $B_N N^{-7/12}$  is bounded in probability.  $\blacksquare$

Note that the function  $s$  in Theorem 3.4 satisfies (3.1), with

$$i(t) = - \int_0^\infty s'(t-u) e^{-\mu u} du = - \int_{-\infty}^t s'(v) e^{-\mu(t-v)} dv. \quad (3.4)$$

This is because the Laplace transform  $\psi$  of the limiting branching random variable  $\tilde{W}$  satisfies the equation

$$\log \psi(v) = - \int_0^\infty \alpha e^{-\mu u} (1 - \psi(v e^{-\lambda u})) du,$$

and substituting  $s(t) = \psi(e^{(\alpha-\mu)t})$  and differentiating gives  $s'(t)/s(t) = -\alpha i(t)$ . Then, differentiating the final expression in (3.4) gives  $i'(t) = -s'(t) - \mu i(t)$ , and the previous

equation for  $s'(t)$  completes (3.1). However, the explicit expression  $\psi(e^{(\alpha-\mu)t})$  also implies the initial condition — for instance, setting  $t = 0$  gives values for  $s(0)$  and  $i(0)$  expressed in terms of  $\mathbb{E}\{e^{-\widetilde{W}}\}$  and  $\mathbb{E}\{\widetilde{W}e^{-\widetilde{W}}\}$ . Alternatively, for an initial condition at  $-\infty$ , one has  $s(-\infty) = 1$  and  $\lim_{u \rightarrow -\infty} e^{-(\alpha-\mu)u} s'(u) = (\alpha - \mu) \lim_{u \rightarrow -\infty} \psi'(e^{(\alpha-\mu)u}) = -(\alpha - \mu) \mathbb{E}\widetilde{W}$ .

## Acknowledgement

ADB thanks the Department of Statistics and Applied Probability at the National University of Singapore for their kind hospitality while this work was undertaken, and a referee for helpful comments.

## References

- [1] D. J. ALDOUS (2013) When knowing early matters: gossip, percolation and Nash equilibria. In: *Prokhorov and Contemporary Probability Theory*, Eds. A.N. Shiryaev, S.R.S. Varadhan and E.L. Presman, pp. 3–28. Springer Proceedings in Mathematics & Statistics **33**, Springer Verlag, Heidelberg.
- [2] N. T. J. BAILEY (1953) The total size of a general stochastic epidemic. *Biometrika* **40**, 177–185.
- [3] F. BALL (1983) The threshold behaviour of epidemic models. *J. Appl. Prob.* **20**, 227–241.
- [4] F. BALL, D. MOLLISON & G. SCALIA–TOMBA (1997) Epidemics with two levels of mixing. *Ann. Appl. Prob.* **7**, 46–89.
- [5] A. D. BARBOUR (1975) The duration of the closed stochastic epidemic. *Biometrika* **62**, 477–482.
- [6] A. D. BARBOUR (2010) Coupling a branching process to an infinite dimensional epidemic process. *ESAIM: P&S* **14**, 53–64.
- [7] A. D. BARBOUR, L. HOLST & S. JANSON (1992) *Poisson approximation*. Clarendon Press, Oxford.
- [8] A. D. BARBOUR & G. REINERT (2013a) Asymptotic behaviour of gossip processes and small world networks. *Adv. Appl. Prob.* **45**, 981–1010.
- [9] A. D. BARBOUR & G. REINERT (2013b) Approximating the epidemic curve. *Electr. J. Prob.* **18**, no. 54, 30pp.
- [10] A. D. BARBOUR & S. UTEV (2004) Approximating the Reed-Frost epidemic process. *Stoch. Procs Applics* **113**, 173–197.
- [11] M. S. BARTLETT (1949) Some evolutionary stochastic processes. *J. Roy. Statist. Soc.* **11**, 211–229.

- [12] S. BHAMIDI, R. VAN DER HOFSTAD & G. HOOGHIEMSTRA (2012) Universality for first passage percolation on sparse random graphs. [arXiv:1210.6839](https://arxiv.org/abs/1210.6839).
- [13] S. CHATTERJEE & R. DURRETT (2011) Asymptotic behavior of Aldous' gossip process. *Ann. Appl. Prob.* **21**, 2447–2482.
- [14] M. GREENWOOD (1931) On the statistical measure of infectiousness. *J. Hygiene* **31**, 336–351.
- [15] D. G. KENDALL (1956) Deterministic and stochastic epidemics in closed populations. *Proc. Third Berkeley Symp. Math. Statist. Prob.* **4**, 149–165.
- [16] W. O. KERMACK & A. G. MCKENDRICK (1927, 1991) Contributions to the mathematical theory of epidemics, part I. *Proc. Roy. Soc. Edin. A* **115**, 700–721; *Bull. Math. Biol.* **53**, 33–55.
- [17] A. G. MCKENDRICK (1925) Applications of mathematics to medical problems. *Proc. Edinburgh Math. Soc.*, **44**, 1–34.
- [18] T. SELLEKE (1983) On the asymptotic distribution of the size of a stochastic epidemic. *J. Appl. Prob.* **20**, 390–394.
- [19] P. WHITTLE (1955) The outcome of a stochastic epidemic — A note on Bailey's paper. *Biometrika* **42**, 116–122.