J. Appl. Prob. Spec. Vol. **51A,** 43–56 (2014)

CELEBRATING 50 YEARS OF THE APPLIED PROBABILITY TRUST

Edited by
S. ASMUSSEN, P. JAGERS, I. MOLCHANOV and L. C. G. ROGERS

Part 3. Biological applications

COUPLINGS FOR LOCALLY BRANCHING EPIDEMIC PROCESSES

A. D. BARBOUR, *Universität Zürich and National University of Singapore*
Institut für Mathematik, Universität Zürich, Winterthurertrasse 190, CH-8057 Zürich.
Email address: a.d.barbour@math.uzh.ch

APPLIED PROBABILITY TRUST
DECEMBER 2014

https://doi.org/10.1239/jap/1417528465 Published online by Cambridge University Press

COUPLINGS FOR LOCALLY BRANCHING EPIDEMIC PROCESSES

BY A. D. BARBOUR

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 Whittle's (1955) branching approximation to the early stages of the epidemic, the phase in which approximately exponential growth takes place. The second is the susceptibility approximation; the backward branching process that approximates the history of the contacts that would lead to an individual becoming infected. The simplest coupling arguments for demonstrating the closeness of these branching process approximations do not keep the processes identical for quite long enough. Thus, arguments showing that the differences are unimportant are also needed. In this paper we show that, for 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

2010 Mathematics Subject Classification: Primary 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 1926 paper in the *Proceedings of the Edinburgh Mathematical Society* 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 Series 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.

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 and Durrett (2011) studied the Aldous (2013) 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

https://doi.org/10.1239/jap/1417528465 Published online by Cambridge University Press

followed by an almost deterministic progression. Barbour and 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 and 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 are 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 the contact probability (homogeneous mixing), and, if t_N is correctly specified, this probability is asymptotically nontrivial. 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 λ 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 $c_{\mathrm{f}} W_{\mathrm{f}}\sqrt{N}$ 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 $C_{\mathrm{b}} W_{\mathrm{b}}\sqrt{N}\mathrm{e}^{\lambda s}$; 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 $1 - \exp(-c_{\mathrm{f}} C_{\mathrm{b}} W_{\mathrm{f}} W_{\mathrm{b}}\mathrm{e}^{\lambda s})$. 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

$$\mathbb{E}[\mathrm{e}^{-c_{\mathrm{f}} C_{\mathrm{b}} W_{\mathrm{f}} W_{\mathrm{b}}\mathrm{e}^{\lambda s}} \mid W_{\mathrm{f}}] = \psi_B(\mathrm{e}^{\lambda s + \log(c_{\mathrm{f}} C_{\mathrm{b}}) + \log W_{\mathrm{f}}}) = s_B(s + \lambda^{-1}[\log(c_{\mathrm{f}} C_{\mathrm{b}}) + \log W_{\mathrm{f}}]),$$

where $\psi_B(v) := \mathbb{E}[\mathrm{e}^{-vW_{\mathrm{b}}}]$ and $s_B(u) := \psi_B(\mathrm{e}^{\lambda u})$. 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 can be shown to be close to $[s_B(s+\lambda^{-1}[\log(c_{\mathrm{f}} C_{\mathrm{b}}) + \log W_{\mathrm{f}}])]^2$, 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 $c_{\mathrm{f}} C_{\mathrm{b}}$ 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

https://doi.org/10.1239/jap/1417528465 Published online by Cambridge University Press

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 and Utev (2004) and 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

$$\mathcal{L}(I_{n+1} \mid S_n, I_n) \,=\, \mathrm{bin}(S_n, 1-q^{I_n}), \qquad S_{n+1} = S_n - I_{n+1}. \tag{2.1}$$

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 $j \ge 0$; the epidemic has terminated. Note also that $S_n = S_0 - \sum_{l=1}^{n} I_l$, so the path of the process is determined by the values of $\{I_n,\ n \ge 0\}$. 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 2.1 below, shows that, with high probability when N is large, the proportion of susceptibles $N^{-1}S_{R(N)+r}$ for all $r \in \mathbb{Z}$ stays close to the points $s(u_N + V_N + r)$ on a fixed continuous, decreasing curve s. 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 $N \to \infty$ to a nontrivial limit V. The distribution of V and the function s can both be deduced from the branching process $Z = \{Z_n,\ n \ge 0\}$ with offspring distribution $\mathrm{Poi}(m)$. 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

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

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 $M \ge s$ and i in $\mathbb{Z}_+$ and for $0 \le p = 1-q \le 1$,*

$$d_{\mathrm{TV}}(\mathrm{bin}(s, 1-q^i), \mathrm{Poi}(Mip)) \,\le\, ip + (Mip)^{-1/2}\big[(M-s)ip + \tfrac{1}{2}si(i-1)p^2\big].$$

If also $M - s \ge i$ then

$$d_{\mathrm{TV}}(\mathrm{bin}(s, 1-q^i), \mathrm{Poi}(Mip)) \,\le\, \frac{1}{2}(M-s)\sqrt{\frac{ip}{M}}\,(3+2Mp).$$

https://doi.org/10.1239/jap/1417528465 Published online by Cambridge University Press

Proof. Since, by Taylor's theorem, $(1-q^i) \le ip$ and $0 \le ip - 1 + q^i \le \frac{1}{2}i(i-1)p^2$, it follows from Barbour *et al.* (1992, Chapter 1, Equation (1.23) and Theorem 1.C(i)) that

$$d_{\mathrm{TV}}(\mathrm{bin}(s, 1-q^i), \mathrm{Poi}(s(1-q^i))) \le ip$$

and that

$$d_{\mathrm{TV}}(\mathrm{Poi}(s(1-q^i)), \mathrm{Poi}(Mip)) \le (Mip)^{-1/2}\big[(M-s)ip + \tfrac{1}{2}si(i-1)p^2\big].$$

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

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

and $s(i-1)p/(M-s) \le sp \le Mp$, from which the last part follows.

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

$$\mathcal{L}(Z_{n+1} \mid Z_n) = \mathrm{Poi}(MpZ_n) \tag{2.3}$$

for the branching process.

Lemma 2.2. *If $M \ge S_0$ then the Reed–Frost process of (2.1) with $S_0 + I_0 = N$ and the branching process Z with offspring distribution* $\mathrm{Poi}(Mp)$*, both starting with $Z_0 = I_0$, can be coupled so that $Z_n = I_n$, $0 \le n \le n_0$, 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 $x \ge 1+\delta$ for $\delta > 0$. In particular, if $m > 1$ and $\alpha > 0$ are fixed, $p = m/(N-1)$, $n_0 \le \alpha \log N / \log m$, and $M - S_0 \le N^\alpha$, then, as $N \to \infty$, the probability of failure equals $O(N^{-1+3\alpha/2})$, and, hence, is small for any $\alpha < \frac{2}{3}$ 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 $0 \le l \le n$, it follows that $S_n = N - B_n \ge S_0 - B_n$ and that $I_n = Z_n$. Comparing recursions (2.1) and (2.3), noting that $M - S_n \ge I_n$ and using Lemma 2.1, it follows that Z_{n+1} and I_{n+1} can be coupled in such a way that

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

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

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

It thus follows that $\mathbb{E}[\sqrt{Z_n}\,] \le \widetilde{m}^{n/2}\sqrt{I_0}$ and that

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

https://doi.org/10.1239/jap/1417528465 Published online by Cambridge University Press

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

$$c(\widetilde{m}) I_0^{3/2} M^{-1} \widetilde{m}^{n_0/2} [(M - S_0) + \widetilde{m}^{n_0}]$$

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

Observe 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 $0 < \alpha < \frac{2}{3}$, let $n_\alpha(N, I_0)$ be the integer such that

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

Define $n_{\mathrm{f}}(N) := n_{7/12}(N, I_0)$ and $n_{\mathrm{b}}(N) := n_{5/12}(N, 1)$, and correspondingly $\theta_{\mathrm{f}}(N)$ and $\theta_{\mathrm{b}}(N)$. Note that $n_{\mathrm{b}}(N)$ 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 offspring distribution Poi(m) can be coupled exactly to the process $\{I_n\}$ of infectives for $n_{\mathrm{f}}(N)$ steps with failure probability at most equal to $O(N^{-1/8})$.

For the main theorem of this section, we condition on $\{Z_n : n = 0, 1, \dots, n_{\mathrm{f}}(N)\}$, denoting this information by $\mathcal{F}_0$. We establish the almost deterministic development from time $n_{\mathrm{f}}(N)$ onwards only on an event in $\mathcal{F}_0$ of asymptotically high probability. Let $\mathcal{C}_N$ denote the event that the coupling up to time $n_{\mathrm{f}}(N)$ is successful, and set $A_N^{(1)} := \{B_{n_{\mathrm{f}}(N)} \le N^{2/3}\}$ and $A_N^{(2)} := \{|\log W_{n_{\mathrm{f}}(N)}| \le \frac{1}{24} \log N\}$, where, as before, $W_n := m^{-n} Z_n$ and $W := \lim_{n\to\infty} W_n$. 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 nonextinction 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 2.1. *For the Reed–Frost epidemic process defined above and any $\varepsilon > 0$, as $N \to \infty$,*

$$\mathbb{P}\left\{ \sup_{r \in \mathbb{Z}} \left| N^{-1} S_{n_{\mathrm{f}}(N) + n_{\mathrm{b}}(N) + r} - s\left(r + \frac{\log c_N + \log W_{n_{\mathrm{f}}(N)}}{\log m} \right) \right| > \varepsilon \;\middle|\; \mathcal{F}_0 \cap A_N^* \right\} \to 0,$$

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

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

$$\begin{aligned} &\mathbb{E}[N^{-1} S_{n_{\mathrm{f}}(N) + n_{\mathrm{b}}(N) + r} \mid \mathcal{F}_0 \cap \mathcal{C}_N] \\ &\quad = \mathbb{P}\{K \in \mathcal{S}_{n_{\mathrm{f}}(N) + n_{\mathrm{b}}(N) + r} \mid \mathcal{F}_0 \cap \mathcal{C}_N\} \\ &\quad = \mathbb{P}\{K \notin \mathcal{B}_{n_{\mathrm{f}}(N) - 1}\} \, \mathbb{P}\{\widetilde{\mathcal{U}}_{n_{\mathrm{b}}(N) + r} \cap \mathcal{Z}_{n_{\mathrm{f}}(N)} = \varnothing \mid \mathcal{F}_0 \cap \mathcal{C}_N \cap \{K \notin \mathcal{B}_{n_{\mathrm{f}}(N) - 1}\}\}, \end{aligned} \tag{2.6}$$

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

$$\begin{aligned} &\mathbb{P}\{\widetilde{\mathcal{U}}_{n_{\mathrm{b}}(N)+r} \cap \mathcal{Z}_{n_{\mathrm{f}}(N)} = \varnothing \mid \mathcal{F}_0 \cap \mathcal{C}_N \cap \{K \notin \mathcal{B}_{n_{\mathrm{f}}(N)-1}\};\ \widetilde{U}_{n_{\mathrm{b}}(N)+r}\} \\ &\quad = P(N - B_{n_{\mathrm{f}}(N)-1}, Z_{n_{\mathrm{f}}(N)}, \widetilde{U}_{n_{\mathrm{b}}(N)+r}), \end{aligned} \tag{2.7}$$

https://doi.org/10.1239/jap/1417528465 Published online by Cambridge University Press

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 $\widetilde{U}_l$, $l \geq 0$, again arise as a Reed–Frost epidemic, this time on the set $[N] \setminus \mathcal{B}_{n_{\mathrm{f}}(N)-1}$ 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 $(\widetilde{Z}_l,\ l \geq 0)$ with $\widetilde{Z}_0 = 1$ and with offspring distribution $\mathrm{Poi}(m)$, with exact coupling failing to hold up to time $n_{\mathrm{b}}(N) + r$ 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 $\widetilde{B}_l$ to denote $\sum_{j=0}^{l} \widetilde{Z}_j$ and $\widetilde{W} := \lim_{n\to\infty} m^{-n}\widetilde{Z}_n$. Noting also that $\mathbb{P}\{K \in \mathcal{B}_{n_{\mathrm{f}}(N)-1}\} \leq N^{-1/3}$ on $A_N^{(1)}$, it follows from (2.6), and by taking the expectation with respect to $\widetilde{U}_{n_{\mathrm{b}}(N)+r}$ in (2.7), that

$$\begin{aligned}\mathbb{E}[N^{-1}S_{n_{\mathrm{f}}(N)+n_{\mathrm{b}}(N)+r} \mid \mathcal{F}_0 \cap \mathcal{C}_N \cap A_N^{(1)}] \\ = \mathbb{E}[P(N - B_{n_{\mathrm{f}}(N)-1}, Z_{n_{\mathrm{f}}(N)}, \widetilde{B}_{n_{\mathrm{b}}(N)+r}) \mid \mathcal{F}_0 \cap \mathcal{C}_N \cap A_N^{(1)}] + O(N^{-1/12}).\end{aligned} \tag{2.8}$$

Furthermore,

$$0 \leq \exp\left(-\frac{m_1 m_2}{n}\right) - P(n, m_1, m_2) \leq n^{-1}(m_1 + m_2),$$

by Barbour *et al.* (1992, Theorem 6A), and, for $0 \leq x \leq m$ and $0 \leq y \leq n/2$,

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

Hence,

$$\begin{aligned}\left|\mathbb{E}\left[\frac{S_{n_{\mathrm{f}}(N)+n_{\mathrm{b}}(N)+r}}{N} \;\middle|\; \mathcal{F}_0 \cap \mathcal{C}_N \cap A_N^{(1)}\right] - \mathbb{E}\left[\exp\left(-\frac{Z_{n_{\mathrm{f}}(N)}\widetilde{B}_{n_{\mathrm{b}}(N)+r}}{N}\right) \;\middle|\; \mathcal{F}_0 \cap \mathcal{C}_N\right]\right| \\ \leq C'(N^{-1/3} + N^{-7/12} + N^{-1/12})\end{aligned} \tag{2.9}$$

for a constant C'.

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

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

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

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

so that, in the same range of r,

$$\mathbb{E}\left|m^{-(n_{\mathrm{b}}(N)+r)}\left[\widetilde{B}_{n_{\mathrm{b}}(N)+r} - \frac{m\widetilde{Z}_{n_{\mathrm{b}}(N)+r}}{m-1}\right]\right| = O(N^{-1/6}\log N).$$

https://doi.org/10.1239/jap/1417528465 Published online by Cambridge University Press

Thus, we conclude that, for $|r \log m| \le \frac{1}{12} \log N$,

$$\left| \mathbb{E}\left[\exp\left(-\frac{Z_{n_{\mathrm{f}}(N)} \widetilde{B}_{n_{\mathrm{b}}(N)+r}}{N} \right) - \exp\left(-\frac{m^{r+1} W_{n_{\mathrm{f}}(N)} \widetilde{W}}{(m-1) I_0 \theta_{\mathrm{f}}(N) \theta_{\mathrm{b}}(N)} \right) \,\middle|\, \mathcal{F}_0 \cap \mathcal{C}_N \right] \right| = O(N^{-1/12}), \tag{2.10}$$

since also, from (2.5), $N = I_0 \theta_{\mathrm{f}}(N) \theta_{\mathrm{b}}(N) m^{n_{\mathrm{f}}(N)+n_{\mathrm{b}}(N)}$. Combining (2.9) and (2.10), it follows that

$$\mathbb{E}\left[\frac{S_{n_{\mathrm{f}}(N)+n_{\mathrm{b}}(N)+r}}{N} \,\middle|\, \mathcal{F}_0 \cap \mathcal{C}_N \cap A_N^{(1)} \right] = \psi(c_N W_{n_{\mathrm{f}}(N)} m^r) + O(N^{-1/12}),$$

where $c_N := m/[(m-1) I_0 \theta_{\mathrm{f}}(N) \theta_{\mathrm{b}}(N)]$, uniformly in $|r \log m| \le \frac{1}{12} \log N$, and $W_{n_{\mathrm{f}}(N)}$ can be replaced by W using (2.4) once more, without change to the order of the error.

It remains to examine $\mathrm{var}(N^{-1} S_{n_{\mathrm{f}}(N)+n_{\mathrm{b}}(N)+r} \mid \mathcal{F}_0 \cap \mathcal{C}_N \cap A_N^{(1)})$. For two points K and K' chosen independently and uniformly from $[N]$, we can write, as above,

$$\begin{aligned} &\mathbb{E}\left[\left(\frac{S_{n_{\mathrm{f}}(N)+n_{\mathrm{b}}(N)+r}}{N} \right)^2 \,\middle|\, \mathcal{F}_0 \cap \mathcal{C}_N \cap A_N^{(1)} \right] \\ &\quad = \mathbb{P}\{\{K, K'\} \cap \mathcal{S}_{n_{\mathrm{f}}(N)+n_{\mathrm{b}}(N)+r} = \varnothing \mid \mathcal{F}_0 \cap \mathcal{C}_N \cap A_N^{(1)}\}. \end{aligned} \tag{2.11}$$

Again, only the event $E_N^{(1)} := \{\{K, K'\} \cap \mathcal{B}_{n_{\mathrm{f}}(N)-1} = \varnothing\}$ is important because the complementary probability is of order $O(N^{-1/3})$ on $A_N^{(1)}$. Then, the corresponding sets $\widetilde{\mathcal{U}}^{(i)}_{n_{\mathrm{b}}(N)+r}$, $i = 1, 2$, can be constructed by using independent branching processes $\widetilde{Z}^{(i)}$, with an error of order $O(N^{-1/12})$. Note that the probability of the event $(E_N^{(2)})^c$ that $\widetilde{\mathcal{U}}^{(1)}_{n_{\mathrm{b}}(N)+r}$ and $\widetilde{\mathcal{U}}^{(2)}_{n_{\mathrm{b}}(N)+r}$ intersect, conditional on their sizes $\widetilde{U}^{(i)}_{n_{\mathrm{b}}(N)+r}$ being known, is the hypergeometric probability

$$1 - P(N - B_{n_{\mathrm{f}}(N)-1}, \widetilde{U}^{(1)}_{n_{\mathrm{b}}(N)+r}, \widetilde{U}^{(2)}_{n_{\mathrm{b}}(N)+r}) = O(N^{-1} \widetilde{U}^{(1)}_{n_{\mathrm{b}}(N)+r} \widetilde{U}^{(2)}_{n_{\mathrm{b}}(N)+r})$$

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

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

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

$$\begin{aligned} &\mathbb{P}\{\{K, K'\} \cap \mathcal{S}_{n_{\mathrm{f}}(N)+n_{\mathrm{b}}(N)+r} = \varnothing \mid \mathcal{F}_0 \cap \mathcal{C}_N \cap A_N^{(1)} \cap E_N^{(1)} \cap E_N^{(2)};\ \widetilde{U}^{(i)}_{n_{\mathrm{b}}(N)+r},\ i = 1, 2\} \\ &\quad = P(N - B_{n_{\mathrm{f}}(N)-1}, Z_{n_{\mathrm{f}}(N)}, \widetilde{U}^{(1)}_{n_{\mathrm{b}}(N)+r})\, P(N - B_{n_{\mathrm{f}}(N)-1} - \widetilde{U}^{(1)}_{n_{\mathrm{b}}(N)+r}, Z_{n_{\mathrm{f}}(N)}, \widetilde{U}^{(2)}_{n_{\mathrm{b}}(N)+r}), \end{aligned}$$

and, on $A_N^{(1)}$ and if also $\widetilde{U}^{(1)}_{n_{\mathrm{b}}(N)+r} \le N^{2/3}$, this is equal to

$$\begin{aligned} &P(N - B_{n_{\mathrm{f}}(N)-1}, Z_{n_{\mathrm{f}}(N)}, \widetilde{U}^{(1)}_{n_{\mathrm{b}}(N)+r})\, P(N - B_{n_{\mathrm{f}}(N)-1}, Z_{n_{\mathrm{f}}(N)}, \widetilde{U}^{(2)}_{n_{\mathrm{b}}(N)+r}) \\ &+ O(N^{-2}(\widetilde{U}^{(1)}_{n_{\mathrm{b}}(N)+r} \widetilde{U}^{(2)}_{n_{\mathrm{b}}(N)+r} Z_{n_{\mathrm{f}}(N)})), \end{aligned}$$

since $|P(N, r, s) - P(N - K, r, s)| \le rsK/[N(N-K)]$. Replacing $\widetilde{U}^{(i)}_{n_{\mathrm{b}}(N)+r}$ by $\widetilde{B}^{(i)}_{n_{\mathrm{b}}(N)+r}$ and taking expectations, and noting the independence of $\widetilde{B}^{(1)}$ and $\widetilde{B}^{(2)}$, this gives

$$\begin{aligned} &\mathbb{P}\{\{K, K'\} \cap \mathcal{S}_{n_{\mathrm{f}}(N)+n_{\mathrm{b}}(N)+r} = \varnothing \mid \mathcal{F}_0 \cap \mathcal{C}_N \cap A_N^{(1)} \cap E_N^{(1)}\} \\ &\quad = (\mathbb{E}[P(N - B_{n_{\mathrm{f}}(N)-1}, Z_{n_{\mathrm{f}}(N)}, \widetilde{B}^{(1)}_{n_{\mathrm{b}}(N)+r}) \mid \mathcal{F}_0 \cap \mathcal{C}_N])^2 + O(N^{-1/12}), \end{aligned}$$

https://doi.org/10.1239/jap/1417528465 Published online by Cambridge University Press

which in turn implies the same estimate for the probability without conditioning on $E_N^{(1)}$. In view of (2.11) and (2.8), this shows that $\mathrm{var}(N^{-1}S_{n_{\mathrm{f}}(N)+n_{\mathrm{b}}(N)+r} \mid \mathcal{F}_0 \cap \mathcal{C}_N \cap A_N^{(1)})$ 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 $\widetilde{Z}$, 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 $W_{n_{\mathrm{f}}(N)}$, there are no more than $(b_\varepsilon - a_\varepsilon) + 3$ values of r such that

$$r + \frac{\log(c_N) + \log W_{n_{\mathrm{f}}(N)}}{\log m} \in (a_\varepsilon - 1, b_\varepsilon + 1).$$

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

The statement of the theorem is complicated by the fact that the process is significantly discrete in the range in which an 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

$$s_n - s_{n+1} \approx s_n\left(1 - \left(1 - \frac{m}{N}\right)^{N(s_{n-1}-s_n)}\right) \approx s_n(1 - \mathrm{e}^{-m(s_{n-1}-s_n)}). \tag{2.12}$$

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

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

and this gives

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

which, with $s(r+c)$ replaced by s_r, recovers (2.12) 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 $\{(S_N(t), I_N(t)),\ t \ge 0\}$ on $(\mathbb{Z}_+)^2$, starting with $(S_N(0), I_N(0)) = (N, I_0)$, and having transition rates, for $(S, I) \in (\mathbb{Z}_+)^2$,

$$(S, I) \to (S-1, I+1) \quad \text{at rate } \alpha I(S/N), \qquad (S, I) \to (S, I-1) \quad \text{at rate } \mu I,$$

where α is the *per capita* contact rate of an infective, and μ 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

https://doi.org/10.1239/jap/1417528465 Published online by Cambridge University Press

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^γ 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

$$\frac{\mathrm{d}s}{\mathrm{d}t} = -\alpha s(t) i(t), \qquad \frac{\mathrm{d}i}{\mathrm{d}t} = (\alpha s(t) - \mu) i(t) \tag{3.1}$$

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 $\gamma = \frac{7}{12}$; the formal statement is given in Theorem 3.2 below. 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^γ infectives, where $\frac{1}{2} < \gamma < \frac{2}{3}$. A more detailed description of the final phase, when $I_N(t)$ falls below N^γ 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

$$(S, I) \to (S-1, I+1) \quad \text{at rate } \alpha I, \qquad (S, I) \to (S, I-1) \quad \text{at rate } \mu I,$$

for $(S, I) \in \mathbb{Z} \times \mathbb{Z}_+$, 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^*. Here I^* itself is a linear birth-and-death process with constant *per capita* birth and death rates α and μ.

Theorem 3.1. *Let $\tau_N^\gamma := \inf\{t \geq 0 \colon I_N(t) \geq N^\gamma\}$, 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, $\lim_{N\to\infty} p_N^C(\gamma) = 1$ *if* $\gamma < \frac{2}{3}$.

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 mth 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 $Z_m \sim \mathrm{bin}(m, p)$ and $p := \alpha/(\alpha+\mu) > \frac{1}{2}$. 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. Denote the set of individuals born up to time τ_N^γ in the epidemic by $\mathcal{B}_N$, for the fixed value $\gamma = \gamma_0 := \frac{7}{12}$, and write $B_N := |\mathcal{B}_N|$, which is also equal to $B^*(\tau_N^{\gamma_0})$ when 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

https://doi.org/10.1239/jap/1417528465 Published online by Cambridge University Press

that a contact would take place before the end of K''s infectious period, when the contact rate is α/N and the infectious period has the Exp(μ) distribution. Conditional on this unlikely event, the distribution of the time to infection is 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 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 $\rho_N\alpha\,\mathrm{e}^{-(\mu+\alpha/N)u}$ at $u\in\mathbb{R}_+$ 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 $(n+1)$th 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 Exp(μ)-distributed infectious periods were shorter than the independent and Exp(α/N)-distributed times to the first points in the infection processes along the links from them to individuals $K_0, K_1, \ldots, K_{n-1}$. 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 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 $\rho_N\alpha\,\mathrm{e}^{-(\mu+(n+1)\alpha/N)u}$ on $\mathbb{R}_+$. We now show that a Poisson process with this intensity can be replaced by one with intensity $\rho_N\alpha\,\mathrm{e}^{-\mu u}$, 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 $\alpha\,\mathrm{e}^{-\mu u}\,\mathrm{d}u$ 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 3.1. *For fixed $\alpha, \mu > 0$, let Π and Ξ be Poisson processes with intensities $\alpha\,\mathrm{e}^{-\mu u}$ and $\alpha\,\mathrm{e}^{-\nu u}$, respectively. Then, uniformly in $\nu \geq \mu$,*

$$d_{\mathrm{TV}}(P,Q) \;=\; O\biggl(\Bigl|1-\frac{\mu}{\nu}\Bigr|\biggr),$$

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

Proof. The processes Π and Ξ each have almost surely only finitely many points. The likelihood of a realization of Π with k points at times $t_1 < t_2 < \cdots < t_k < \infty$ is

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

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

$$\frac{\mathrm{d}Q}{\mathrm{d}P}(t_1,\ldots,t_k) \;=\; \mathrm{e}^{-(\mu-\nu)[\alpha/(\mu\nu)-s_k]}.$$

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

$$\mathbb{E}_P[\mathrm{e}^{\theta S}] \;=\; \exp\biggl(\alpha\int_0^\infty \mathrm{e}^{-\mu u}(\mathrm{e}^{\theta u}-1)\,\mathrm{d}u\biggr) \;=\; \exp\biggl(\frac{\alpha\theta}{\mu(\mu-\theta)}\biggr)$$

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

$$\mathbb{E}_P\biggl[\biggl(\frac{\mathrm{d}Q}{\mathrm{d}P}(\{T_j,\ j\geq 1\})\biggr)^2\biggr] \;=\; \exp\biggl(\frac{2\alpha(\nu-\mu)}{\mu\nu}\biggr)\mathbb{E}_P[\mathrm{e}^{-2(\nu-\mu)S}] \;=\; \exp\biggl(\frac{2\alpha(\nu-\mu)^2}{\mu\nu(2\nu-\mu)}\biggr).$$

https://doi.org/10.1239/jap/1417528465 Published online by Cambridge University Press

It is then immediate that $\mathrm{var}_P\{\mathrm{d}Q/\mathrm{d}P\} = O((1-\mu/\nu)^2)$, 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

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

Lemma 3.2. *Let Z_a denote a Crump–Mode–Jagers branching process with $Z_a(0) = 1$, whose offspring point processes are distributed as Poisson processes P_a with intensities $a\,\mathrm{e}^{-\mu u}$ at $u \in \mathbb{R}_+$. Then, for $\alpha_1 < \alpha_2 < 3\alpha_1/2$ and any $m \geq 1$,*

$$d_{\mathrm{TV}}(\mathcal{L}(\{Z_{\alpha_1}(s),\ 0 \leq s \leq \tau^{(1)}_{m+1}\}), \mathcal{L}(\{Z_{\alpha_2}(s),\ 0 \leq s \leq \tau^{(2)}_{m+1}\})) \;\leq\; \sqrt{\frac{9me}{8}}\left|1 - \frac{\alpha_1}{\alpha_2}\right|,$$

where $\tau^{(l)}_m$ denotes the birth time of the mth individual in Z_{α_l}.

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

$$1 \;\leq\; \mathbb{E}_{Q_b}\left[\left(\frac{\mathrm{d}Q_a}{\mathrm{d}Q_b}(U)\right)^2\right] \;=\; 1 + \frac{(a-b)^2}{b(2a-b)}\{1 - \mathrm{e}^{-(2a-b)F(\infty)}\} \;\leq\; 1 + \frac{(a-b)^2}{b(2a-b)}.$$

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

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

Hence, $\mathbb{E}[W_{j+1} \mid \mathcal{F}_j] = W_j$ and

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

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

$$\mathrm{var}_{\alpha_1} W_{j+1} \;\leq\; \mathrm{var}_{\alpha_1} W_j(1+\varepsilon) + \varepsilon \;\leq\; j\varepsilon\mathrm{e}^{j\varepsilon},$$

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

As a result of Lemmas 3.1 and 3.2, and because $\mathbb{E}U_r \leq r\alpha/\mu$, it follows that the susceptibility process from any K', chosen uniformly at random from $[N]\backslash\mathcal{B}_N$, can be replaced by a branching process $\widetilde{Z}$ with the distribution of Z_α as in Lemma 3.2, until the birth of the $(m+1)$th 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_α has Malthusian parameter $\lambda := \alpha - \mu > 0$ and $\mathbb{E}[(\int_0^\infty \mathrm{e}^{-\lambda u}\,P_\alpha(\mathrm{d}u))^2] < \infty$, it follows that

https://doi.org/10.1239/jap/1417528465 Published online by Cambridge University Press

$C := \sup_t \mathbb{E}[Z_\alpha(t)\mathrm{e}^{-(\alpha-\mu)t}] < \infty$. Hence, defining

$$t_N(u) := \frac{1}{\alpha-\mu}\frac{5}{12}\log N + u, \qquad m_N := N^{23/48}, \qquad u_N := \frac{\log N}{24(\alpha-\mu)},$$

it follows by Markov's inequality that

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

Thus, on $\{B_N \le N^{2/3}\}$, the susceptibility process from K' can be coupled to Z_α 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 and Reinert (2013b), because the exponential distribution of an individual's infectious period does not satisfy their Assumption 2; see their Remark 2.1.

Theorem 3.2. *For the Markovian SIR epidemic process as defined above, and any* $\varepsilon > 0$,

$$\lim_{N\to\infty} \mathbb{P}\Big\{\sup_{u\in\mathbb{R}} |N^{-1}S(\tau_N^{\gamma_0}+t_N(u)) - s(u)| > \varepsilon \;\Big|\; \mathcal{F}_0 \cap \{\tau_N^{\gamma_0} < \infty\} \cap \{B_N \le N^{2/3}\}\Big\} = 0,$$

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

Proof. The proof is similar in spirit to that of Theorem 2.1. 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 $\widetilde{Z}$ can be taken to be distributed as Z_α up to time $t_N(u)$, with error probability of order $O(N^{-1/48})$ for any $u \le u_N$, and at time $t = t_N(u)$ it has $\widetilde{Z}_t = \widetilde{W}_t N^{5/12}\mathrm{e}^{(\alpha-\mu)u}$ individuals.

Let the jth of the N_{γ_0} individuals $\mathcal{B}_{\gamma_0}$ of I_N at time $\tau_N^{\gamma_0}$ have C_j direct offspring born at times $\{\tau_{jl},\ 1 \le l \le C_j\}$. 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

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

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

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

because $\mathbb{P}\{A_r \ge v\} = \widetilde{W}_{t_N(u)-v}\mathrm{e}^{-(\alpha-\mu)v}/\widetilde{W}_{t_N(u)} = \mathrm{e}^{-(\alpha-\mu)v}(1+O(N^{-5/24}))$, and

$$\mathbb{E}\Big[\sum_{l=1}^{C_j}\mathrm{e}^{-(\alpha-\mu)\tau_{jl}}\Big] = 1.$$

Thus, $s(u)$ approximates $N^{-1}\mathbb{E}[S(\tau_N^{\gamma_0}+t_N(u)) \mid \mathcal{F}_0 \cap \{\tau_N^{\gamma_0} < \infty\} \cap \{B_N \le N^{2/3}\}]$, and the rest of the proof is then much as for Theorem 2.1. That

$$\lim_{N\to\infty}\mathbb{P}\{\{\tau_N^{\gamma_0} < \infty\} \cap \{B_N > N^{2/3}\}\} = 0$$

https://doi.org/10.1239/jap/1417528465 Published online by Cambridge University Press

follows because, on $W^* > 0$, $B^*(t)/I^*(t) \to \alpha/(\alpha-\mu)$ almost surely, and, hence, $B_N N^{-7/12}$ is bounded in probability.

Note that the function s in Theorem 3.2 satisfies (3.1), with

$$i(t) \;=\; -\int_0^\infty s'(t-u)\mathrm{e}^{-\mu u}\,\mathrm{d}u \;=\; -\int_{-\infty}^t s'(v)\mathrm{e}^{-\mu(t-v)}\,\mathrm{d}v. \tag{3.2}$$

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

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

and substituting $s(t) = \psi(\mathrm{e}^{(\alpha-\mu)t})$ and differentiating gives $s'(t)/s(t) = -\alpha\, i(t)$. Then, differentiating the final expression in (3.2) gives $i'(t) = -s'(t) - \mu i(t)$, and the previous equation for $s'(t)$ completes (3.1). However, the explicit expression $\psi(\mathrm{e}^{(\alpha-\mu)t})$ also implies the initial condition—for instance, setting $t = 0$ gives values for $s(0)$ and $i(0)$ expressed in terms of $\mathbb{E}[\mathrm{e}^{-\widetilde{W}}]$ and $\mathbb{E}[\widetilde{W}\mathrm{e}^{-\widetilde{W}}]$. Alternatively, for an initial condition at $-\infty$, we have $s(-\infty) = 1$ and

$$\lim_{u\to-\infty}\mathrm{e}^{-(\alpha-\mu)u}s'(u) = (\alpha-\mu)\lim_{u\to-\infty}\psi'(\mathrm{e}^{(\alpha-\mu)u}) = -(\alpha-\mu)\mathbb{E}\widetilde{W}.$$

Acknowledgements

ADB was Saw Swee Hock Professor of Statistics at the National University of Singapore while this work was undertaken; the work was also supported in part by Australian Research Council grants DP120102728 and DP120102398. ADB thanks the anonymous referee for helpful comments.

References

Aldous, D. J. (2013). When knowing early matters: gossip, percolation and Nash equilibria. In *Prokhorov and Contemporary Probability Theory* (Proc. Math. Statist. **33**), eds. A. N. Shiryaev, S. R. S. Varadhan and E. L. Presman, Springer, Heidelberg, pp. 3–28.

Bailey, N. T. J. (1953). The total size of a general stochastic epidemic. *Biometrika* **40,** 177–185.

Ball, F. (1983). The threshold behaviour of epidemic models. *J. Appl. Prob.* **20,** 227–241.

Barbour, A. D. (1975). The duration of the closed stochastic epidemic. *Biometrika* **62,** 477–482.

Barbour, A. D. (2010). Coupling a branching process to an infinite dimensional epidemic process. *ESAIM Prob. Statist.* **14,** 53–64.

Barbour, A. D. and Reinert, G. (2013a). Asymptotic behaviour of gossip processes and small-world networks. *Adv. Appl. Prob.* **45**, 981–1010.

Barbour, A. D. and Reinert, G. (2013b). Approximating the epidemic curve. *Electron. J. Prob.* **18,** no. 54, 30pp.

Barbour, A. D. and Utev, S. (2004). Approximating the Reed–Frost epidemic process. *Stoch. Process. Appl.* **113,** 173–197.

Barbour, A. D., Holst, L. and Janson, S. (1992). *Poisson approximation*. Oxford University Press.

Bartlett, M. S. (1949). Some evolutionary stochastic processes. *J. R. Statist. Soc. B* **11,** 211–229.

Bhamidi, S., van der Hofstad, R. and Hooghiemstra, G. (2012). Universality for first passage percolation on sparse random graphs. Preprint. Available at http://arxiv.org/abs/1210.6839v1.

Chatterjee, S. and Durrett, R. (2011). Asymptotic behavior of Aldous' gossip process. *Ann. Appl. Prob.* **21,** 2447–2482.

Greenwood, M. (1931). On the statistical measure of infectiousness. *J. Hygiene* **31,** 336–351.

Kendall, D. G. (1956). Deterministic and stochastic epidemics in closed populations. In *Proc. 3rd Berkeley Symp. Math. Statist. Prob.*, Vol. IV, University of California Press, Berkeley, pp. 149–165.

Kermack, W. O. and McKendrick, A. G. (1927). Contributions to the mathematical theory of epidemics. *Proc. R. Soc. London A* **115,** 700–721.

McKendrick, A. G. (1926). Applications of mathematics to medical problems. *Proc. Edinburgh Math. Soc.* **44,** 98–130.
Sellke, T. (1983). On the asymptotic distribution of the size of a stochastic epidemic. *J. Appl. Prob.* **20,** 390–394.
Whittle, P. (1955). The outcome of a stochastic epidemic—a note on Bailey's paper. *Biometrika* **42,** 116–122.

A. D. BARBOUR, *Universität Zürich and National University of Singapore*
Institut für Mathematik, Universität Zürich, Winterthurertrasse 190, CH-8057 Zürich.
Email address: a.d.barbour@math.uzh.ch