STRONG CONVERGENCE OF AN EPIDEMIC MODEL WITH MIXING GROUPS

FRANK BALL [iD]* ** AND

PETER NEAL [iD],* *** *University of Nottingham*

Abstract

We consider an SIR (susceptible → infective → recovered) epidemic in a closed population of size n, in which infection spreads via mixing events, comprising individuals chosen uniformly at random from the population, which occur at the points of a Poisson process. This contrasts sharply with most epidemic models, in which infection is spread purely by pairwise interaction. A sequence of epidemic processes, indexed by n, and an approximating branching process are constructed on a common probability space via embedded random walks. We show that under suitable conditions the process of infectives in the epidemic process converges almost surely to the branching process. This leads to a threshold theorem for the epidemic process, where a major outbreak is defined as one that infects at least $\log n$ individuals. We show further that there exists $\delta > 0$, depending on the model parameters, such that the probability that a major outbreak has size at least δn tends to one as $n \to \infty$.

Keywords: Branching process; coupling; random walk; SIR epidemic; size of epidemic; threshold theorem

1. Introduction

A key component of any epidemic model is the assumption made concerning transmission of infection between individuals. In almost all epidemic models it is assumed that infection spreads via interactions of *pairs* of individuals, one of whom is infective and the other susceptible. In some epidemic models, such as network models (e.g. Newman [15]), this assumption is explicit, whereas in others, such as the so-called general stochastic epidemic (e.g. Bailey [1, Chapter 6]) and many deterministic models, it is implicit. In the general stochastic epidemic, the process of the numbers of susceptible and infective individuals, $\{(S(t), I(t)) : t \geq 0\}$, is modelled as a continuous-time Markov chain with infinitesimal transition probabilities

$$\mathbb{P}((S(t + \Delta t), I(t + \Delta t)) = (s - 1, i + 1)|(S(t), I(t)) = (s, i)) = \beta si\Delta t + o(\Delta t),$$
$$\mathbb{P}((S(t + \Delta t), I(t + \Delta t)) = (s, i - 1)|(S(t), I(t)) = (s, i)) = \gamma i\Delta t + o(\Delta t),$$

and with all other transitions having probability $o(\Delta t)$. Here, β is the individual-to-individual infection rate and γ is the recovery rate. However, it is probabilistically equivalent to a model

in which the infectious periods of infectives follow independent exponential random variables having mean γ^{-1} and contacts between distinct pairs of individuals occur at the points of independent Poisson processes, each having rate β.

In real-life epidemics, people often meet in groups of size larger than two; in many countries, one of the most significant control measures in the COVID-19 pandemic was to impose limits on the size of gatherings outside of the home. In Cortez [8] and Ball and Neal [5], the authors independently introduced a new class of SIR (susceptible → infective → recovered) epidemic model, in which *mixing events* occur at the points of a Poisson process, with the sizes of successive mixing events being independently distributed according to a random variable having support contained in $\{2, 3, \ldots, n\}$, where n is the population size. Mixing events are instantaneous, and at a mixing event of size c, each infective present contacts each susceptible present independently with probability π_c; a susceptible becomes infected if they are contacted by at least one infective. Such an infected susceptible immediately becomes infective, although they cannot infect other susceptibles at the same mixing event, and remains so for a time that follows an exponential distribution with mean γ^{-1}. In Cortez [8] and Ball and Neal [5], the temporal behaviour of epidemics with many initial infectives is studied, with [8] considering the mean-field limit of the stochastic epidemic process. In Ball and Neal [5], the focus is on a functional central limit theorem for the temporal behaviour of epidemics with many initial infectives and on central limit theorems for the final size of (i) an epidemic with many initial infectives and (ii) an epidemic with few initial infectives that becomes established and leads to a major outbreak. A branching process which approximates the early stages of an epidemic with few initial infectives is described in [5], though no rigorous justification is provided. A key result required in the proof of the central limit theorem for the final size in Case (ii) above is that there exists $\delta > 0$ (which depends on model parameters) such that the probability that an epidemic infects at least a fraction δ of the population, given that it infects at least $\log n$ individuals, converges to one as the population size $n \to \infty$. This result is simply stated without a proof in [5]. The aim of the present paper is to fill these gaps for a model that allows more general transmission of infection at mixing events than that considered in [5].

Approximation of the process of infectives in an epidemic model by a branching process has a long history that goes back to the pioneering work of Bartlett [7, pp. 147–148] and Kendall [12], who considered approximation of the number of infectives in the general stochastic epidemic by a linear birth-and-death process, with birth rate βN and death rate γ, where N is the initial number of susceptibles. This leads to the celebrated threshold theorem (Whittle [17] and Williams [18]), arguably the most important result in mathematical epidemic theory. The approximation was made fully rigorous by Ball [2] (cf. Metz [13]), who defined realisations of the general stochastic epidemic, indexed by N, with the Nth epidemic having infection rate βN^{-1} and recovery rate γ, and the limiting birth-and-death process on a common probability space and used a coupling argument to prove almost sure convergence, as $N \to \infty$, of the epidemic process to the limiting branching process over any finite time interval $[0, t]$. The method was extended by Ball and Donnelly [3] to show almost sure convergence over suitable intervals $[0, t_N]$, where $t_N \to \infty$ as $N \to \infty$.

The key idea of Ball [2] is to construct a realisation of the epidemic process for each N from the same realisation of the limiting branching process. Moreover, this coupling is done on an individual basis, in that the behaviour of an infective in the Nth epidemic model is derived from the behaviour of a corresponding individual in the branching process. The method is very powerful and applicable to a broad range of epidemic models. However, it cannot be easily applied to epidemics with mixing groups, because the mixing groups induce dependencies

between different infectives. Thus instead, we generalise the method of Ball and O'Neill [6], which involves constructing sample paths of the epidemic process, indexed by the population size n, and the limiting branching process (more precisely, the numbers of infectives in the epidemic processes and the number of individuals in the branching process) via a sequence of independent and identically distributed (i.i.d.) random vectors. The generalisation is far from straightforward, since Ball and O'Neill [6] consider only epidemics in which the number of infectives changes in steps of size 1, as in the general stochastic epidemic, whereas in the model with mixing events, although the number of infectives can only decrease in steps of size 1, it can increase in steps of any size not greater than the population size n. We improve on the coupling given in [6] by coupling the time of events in the limiting branching process and epidemic processes, so that the event times agree with high probability, tending to 1 as the population size $n \to \infty$, rather than having the event times in the epidemic processes converge in the limit, as the population size $n \to \infty$, to the event times of the branching process. Finally, we use the coupling to prove the above-mentioned result concerning epidemics of size at least $\log n$, which was not addressed in [6].

The remainder of the paper is structured as follows. The model with mixing groups $\mathcal{E}^{(n)}$ is defined in Section 2. The approximating branching process $\mathcal{B}$ and the main results of the paper are given in Section 3. The branching process $\mathcal{B}$ is described in Section 3.1, where some of its basic properties are presented. The offspring mean of $\mathcal{B}$ yields the basic reproduction number R_0 of the epidemic $\mathcal{E}^{(n)}$. The extinction probability and Malthusian parameter of $\mathcal{B}$ are derived. The main results of the paper are collected together in Section 3.2. Theorem 3.1 shows that the number of infectives in the epidemic process $\mathcal{E}^{(n)}$ converges almost surely to the number of individuals alive in the branching process $\mathcal{B}$ on $[0, t_n)$ as $n \to \infty$, where $t_n = \infty$ in the case the branching process goes extinct and $t_n = \rho \log n$ for some $\rho > 0$ otherwise. A major outbreak is defined as one that infects at least $\log n$ individuals. Theorem 3.2(a) shows that the probability of a major outbreak converges to the survival probability of $\mathcal{B}$ as $n \to \infty$. Theorem 3.2(b) shows that if $R_0 > 1$, so a major outbreak occurs with non-zero probability in the limit $n \to \infty$, then there exists $\delta > 0$ such that the probability that a major outbreak infects at least a fraction δ of the population tends to one as $n \to \infty$. Moreover, we show that there exists $\delta' > 0$ such that the fraction of the population infectious at the peak of the epidemic exceeds δ' with probability tending to one as $n \to \infty$. The proofs of Theorems 3.1 and 3.2 are given in Sections 4 and 5, respectively. Brief concluding comments are given in Section 6.

2. Model

We consider the spread of an SIR epidemic in a closed population of n individuals, with infection spread via mixing events which occur at the points of a homogeneous Poisson process having rate $n\lambda$. The sizes of mixing events are i.i.d. according to a random variable $C^{(n)}$ having support $\{2, 3, \ldots, n\}$. If a mixing event has size c then it is formed by choosing c individuals uniformly at random from the population without replacement. Suppose that a mixing event of size c involves i susceptible and j infective individuals, and hence $c - i - j$ recovered individuals. Then the probability that w new infectives are created at the event is $\pi_c(w; i, j)$. The only restrictions we impose on $\pi_c(w; i, j)$ are the natural ones that, for $w > 0$, $\pi_c(w; i, 0) = 0$; infections can only occur at a mixing event if there is at least one infective present; and for $w > i$, $\pi_c(w; i, j) = 0$: the maximum number of new infectives created at a mixing event is the number of susceptibles involved in the event. Mixing events are assumed to be instantaneous. The infectious periods of infectives follow independent $\text{Exp}(\gamma)$ random variables, i.e. exponential random variables having rate γ and hence mean γ^{-1}. There is no latency period, so

newly infected individuals are immediately able to infect other individuals. (The possibility of their being able to infect other susceptibles during the mixing event at which they were infected can be incorporated into the $\pi_c(w; i, j)$.) All processes and random variables in the above model are mutually independent. The epidemic starts at time $t = 0$ with m_n infective and $n - m_n$ susceptible individuals, and terminates when there is no infective left in the population. Denote this epidemic model by $\mathcal{E}^{(n)}$.

2.1. Special cases

2.1.1. *General stochastic epidemic.* If all mixing groups have size 2, i.e. $\mathbb{P}(C^{(n)} = 2) = 1$, and $\pi_c(0; 1, 1) = \pi_c(1; 1, 1) = \frac{1}{2}$, then the model reduces to the general stochastic epidemic, with individual-to-individual infection rate $\beta = \frac{\lambda}{n-1}$ and recovery rate γ.

2.1.2. *Binomial sampling.* The models studied in Cortez [8] and Ball and Neal [5] make the Reed–Frost-type assumption that at a mixing event of size c, each infective present has probability π_c of making an infectious contact with any given susceptible present, with all such contacts being independent. This corresponds to

$$\pi_c(w; i, j) = \binom{i}{w} \left(1 - (1 - \pi_c)^j\right)^w (1 - \pi_c)^{j(i-w)} \qquad (w = 0, 1, \ldots, i).$$

3. Approximating branching process and main results

3.1. Approximating branching process

We approximate the process of infectives in the early stages of the epidemic $\mathcal{E}^{(n)}$ by a branching process $\mathcal{B}$, which assumes that every mixing event which includes at least one infective consists of a single infective in an otherwise susceptible group. In the epidemic $\mathcal{E}^{(n)}$, the probability that a given mixing event of size c involves a specified individual, i_* say, is $\frac{c}{n}$, so mixing events that include i_* occur at rate

$$\lambda n \sum_{c=2}^{n} p_C^{(n)}(c) \frac{c}{n} = \lambda \mu_C^{(n)}, \tag{3.1}$$

where $p_C^{(n)}(c) = \mathbb{P}(C^{(n)} = c)$ $(c = 2, 3, \ldots, n)$ and $\mu_C^{(n)} = \mathbb{E}[C^{(n)}]$. Furthermore, the probability that a given mixing event is of size c given that it includes i_* is

$$\frac{p_C^{(n)}(c) \frac{c}{n}}{\sum_{c'=2}^{n} p_C^{(n)}(c') \frac{c'}{n}} = \frac{c p_C^{(n)}(c)}{\mu_C^{(n)}} \qquad (c = 2, 3, \ldots, n).$$

Suppose that $C^{(n)} \xrightarrow{\mathrm{D}} C$ as $n \to \infty$ (where $\xrightarrow{\mathrm{D}}$ denotes convergence in distribution), $p_C(c) = \mathbb{P}(C = c)$ $(c = 2, 3, \ldots)$, and $\mu_C^{(n)} \to \mu_C = \sum_{c=2}^{\infty} c p_C(c)$, which we assume to be finite. Thus in the limit as $n \to \infty$, mixing events involving i_* occur at rate $\lambda \mu_C$, and the size of such a mixing event is distributed according to $\tilde{C}$, the size-biased version of C, having probability mass function

$$p_{\tilde{C}}(c) = \mathbb{P}(\tilde{C} = c) = \frac{c p_C(c)}{\mu_C} \qquad (c = 2, 3, \ldots). \tag{3.2}$$

We assume that the initial number of infectives $m_n = m$ for all sufficiently large n, so the branching process $\mathcal{B}$ has m ancestors.

In $\mathcal{B}$, a typical individual, i_* say, has lifetime $L \sim \mathrm{Exp}(\gamma)$, during which they have birth events at rate $\lambda\mu_C$. Let $\tilde{Z}_1, \tilde{Z}_2, \ldots$ denote the number of offspring i_* has at successive birth events. A birth event corresponds to a mixing event involving a single infective in an otherwise susceptible group in the epidemic. Thus, $\tilde{Z}_1, \tilde{Z}_2, \ldots$ are i.i.d. copies of a random variable $\tilde{Z}$, with $\mathbb{P}(\tilde{Z} = w) = \varphi_w$ ($w = 0, 1, \ldots$), where

$$\varphi_w = \sum_{c=w+1}^{\infty} p_{\tilde{C}}(c)\pi_c(w; c-1, 1) = \frac{1}{\mu_C} \sum_{c=w+1}^{\infty} cp_C(c)\pi_c(w; c-1, 1), \tag{3.3}$$

using (3.2). Note that an individual may produce no offspring at a birth event. The number of birth events a typical individual has during their lifetime, G say, has the geometric distribution

$$\mathbb{P}(G = k) = \frac{\gamma}{\gamma + \lambda\mu_C} \left(\frac{\lambda\mu_C}{\gamma + \lambda\mu_C} \right)^k \qquad (k = 0, 1, \ldots). \tag{3.4}$$

Let R be the total number of offspring a typical individual has during their lifetime. Then

$$R = \sum_{i=1}^{G} \tilde{Z}_i, \tag{3.5}$$

where $G, \tilde{Z}_1, \tilde{Z}_2, \ldots$ are independent and the sum is zero if $G = 0$.

The basic reproduction number $R_0 = \mathbb{E}[R]$. Hence, using (3.5) and (3.4),

$$\begin{aligned} R_0 = \mathbb{E}[G]\mathbb{E}[\tilde{Z}] &= \frac{\lambda\mu_C}{\gamma} \sum_{w=1}^{\infty} \frac{w}{\mu_C} \sum_{c=w+1}^{\infty} cp_C(c)\pi_c(w; c-1, 1) \\ &= \frac{\lambda}{\gamma} \sum_{c=2}^{\infty} cp_C(c) \sum_{w=1}^{c-1} w\pi_c(w; c-1, 1) \\ &= \frac{\lambda}{\gamma} \mathbb{E}[Cv(C)], \end{aligned} \tag{3.6}$$

where

$$v(c) = \sum_{w=1}^{c-1} w\pi_c(w; c-1, 1) \tag{3.7}$$

is the mean number of new infectives generated in a mixing event of size c with one infective and $c-1$ susceptibles. Again using (3.5) and (3.4), the offspring probability generating function for the branching process $\mathcal{B}$ is

$$f_R(s) = \mathbb{E}[s^R] = \frac{\gamma}{\gamma + \lambda\mu_C \left(1 - f_{\tilde{Z}}(s)\right)},$$

where $f_{\tilde{Z}}(s) = \sum_{w=0}^{\infty} \varphi_w s^w$. By standard branching process theory, the extinction probability z of $\mathcal{B}$, given that initially there is one individual, is given by the smallest solution in $[0, 1]$ of $f_R(s) = s$. Furthermore, $z < 1$ if and only if $R_0 > 1$.

Let r denote the Malthusian parameter of $\mathcal{B}$; see Jagers [11, p. 10] for details. The mean rate at which an individual produces offspring t time units after their birth is $\mathbb{P}(L > t)\lambda\mu_C\mathbb{E}[\tilde{Z}_1] = \gamma e^{-\gamma t}R_0$ ($t > 0$), so r is the unique solution in $(0, \infty)$ of

$$\int_0^{\infty} e^{-rt}\gamma e^{-\gamma t}R_0 \, dt = 1,$$

whence

$$r = \gamma(R_0 - 1). \tag{3.8}$$

Note that r depends on the parameters of the epidemic model only through (R_0, γ). Thus, if R_0 and γ are held fixed, then the Malthusian parameter is the same for all corresponding choices of the distribution of C and $\{\pi_c(w; i, j)\}$. In particular, under these conditions, the early exponential growth of an epidemic that takes off is the same as that of the general stochastic epidemic.

3.2. Strong convergence of epidemic processes

In this section we consider a sequence of epidemics $(\mathcal{E}^{(n)})$, in which $m_n = m$ for all sufficiently large n, and state results concerned with convergence of the process of infectives in the epidemic process $\mathcal{E}^{(n)}$ to the branching process $\mathcal{B}$ as $n \to \infty$ that are proved in Section 4. The usual approach to proving such results is based upon that of Ball [2] and Ball and Donnelly [3], in which the sample paths of the epidemic process for each n are constructed from those of the limiting branching process, $\mathcal{B}$. As noted in the introduction, that approach is not easily implemented in the present setting, because the mixing groups induce dependencies between different infectives. We therefore generalise the method in Ball and O'Neill [6] and construct sample paths of the epidemic processes and the limiting branching process, $\mathcal{B}$, from a sequence of i.i.d. random vectors defined on an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The construction is described in Section 4.

For $t \geq 0$, let $S^{(n)}(t)$ and $I^{(n)}(t)$ be the numbers of susceptibles and infectives, respectively, at time t in $\mathcal{E}^{(n)}$. Let $T^{(n)} = n - S^{(n)}(\infty)$ be the total size of the epidemic $\mathcal{E}^{(n)}$, i.e. the total number of individuals infected during its course, including the initial infectives. For $t \geq 0$, let $I(t)$ be the number of individuals alive at time t in $\mathcal{B}$, and let T be the total size of the branching process $\mathcal{B}$, including the m ancestors. Note that whereas $T^{(n)}(\omega) < \infty$ for all $\omega \in \Omega$, $T(\omega) = \infty$ if the branching process $\mathcal{B}(\omega)$ does not go extinct.

Throughout the remainder of the paper we assume that $m_n = m$ and $\mu_C^{(n)} \leq \mu_C$ for all sufficiently large n. The assumption $\mu_C^{(n)} \leq \mu_C$ simplifies the presentation of certain results, in particular, Lemma 4.2, and holds in the most common cases: (i) C has finite support $\{2, 3, \ldots, n_0\}$, and for $n \geq n_0$, $C^{(n)} = C$; (ii) $C^{(n)} = \min\{C, n\}$; and (iii) $C^{(n)} \stackrel{D}{=} C|C \leq n$. We also assume throughout that $C^{(n)} \stackrel{D}{\longrightarrow} C$ and $\mathbb{E}[(C^{(n)})^2] \to \mathbb{E}[C^2] < \infty$ as $n \to \infty$. For Theorem 3.1(b) and Theorem 3.2, we require additional conditions on $C^{(n)}$ and C, namely that

$$\lim_{n \to \infty} \mathbb{E}\big[C^{(n)}\big(C^{(n)} - 1\big)\nu\big(C^{(n)}\big)\big] = \mathbb{E}[C(C - 1)\nu(C)] < \infty, \tag{3.9}$$

and that there exists $\theta_0 > 0$ such that

$$\lim_{n \to \infty} n^{\theta_0} \sum_{c=2}^{\infty} c \left| p_C^{(n)}(c) - p_C(c) \right| = 0. \tag{3.10}$$

Note that $\mathbb{E}[\big(C^{(n)}\big)^3] \to \mathbb{E}[C^3] < \infty$ as $n \to \infty$ is a sufficient condition for (3.9) to hold. Also, in the three common cases listed above for constructing $C^{(n)}$ from C, (3.10) holds for any $0 < \theta_0 < \alpha$, for which $\mathbb{E}[C^{1+\alpha}] < \infty$. (For Case (i), this is immediate. For Cases (ii) and (iii), the proof is similar to that of (A1) in the Supplementary Information of Ball and Neal [5].)

Theorem 3.1. *Under the stated conditions on $C^{(n)}$, there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which are defined a sequence of epidemic models, $\mathcal{E}^{(n)}$, indexed by n, and the approximating branching process, $\mathcal{B}$, with the following properties.*

Denote by A_{ext} the set on which the branching process $\mathcal{B}$ becomes extinct:

$$A_{ext} = \{\omega \in \Omega : T(\omega) < \infty\}.$$

(a) *Then, as $n \to \infty$,*

$$\sup_{0 \le t < \infty} \left| I^{(n)}(t) - I(t) \right| \to 0$$

for $\mathbb{P}$-almost all $\omega \in A_{ext}$.

(b) *Suppose that (3.9) holds and (3.10) holds for some $\theta_0 > 0$. Then, if there exists $\alpha \ge 1$ such that $\mathbb{E}[C^{\alpha+1}] < \infty$, we have for*

$$0 < \rho < \frac{1}{r} \min \left\{ \frac{\alpha \theta_0}{2(1+\alpha)}, \frac{\alpha}{2+4\alpha} \right\}, \tag{3.11}$$

as $n \to \infty$,

$$\sup_{0 \le t \le \rho \log n} \left| I^{(n)}(t) - I(t) \right| \to 0 \tag{3.12}$$

for $\mathbb{P}$-almost all $\omega \in A_{ext}^c$.

The proof of Theorem 3.1 is presented in Section 4.

Note that ρ given in (3.11) satisfies $\rho < (4r)^{-1}$, and thus Theorem 3.1(b) is weaker than [3, Theorem 2.1, (2.2)], where (3.12) is shown to hold for $\rho < (2r)^{-1}$ in the standard pairwise mixing epidemic model. The following corollary of Theorem 3.1 concerns the final size of the epidemic.

Corollary 3.1. *For $(\Omega, \mathcal{F}, \mathbb{P})$ defined in Theorem 3.1, we have, for $\mathbb{P}$-almost all $\omega \in \Omega$,*

$$\lim_{n \to \infty} T^{(n)}(\omega) = T(\omega).$$

Corollary 3.1 shows that for large n, the final size of the epidemic $\mathcal{E}^{(n)}$ can be approximated by the total size of $\mathcal{B}$. This leads to a threshold theorem for the epidemic process $\mathcal{E}^{(n)}$ by associating survival (i.e. non-extinction) of the branching process $\mathcal{B}$ with a major outbreak in the epidemic process $\mathcal{E}^{(n)}$ (cf. Ball [2, Theorem 6], and Ball and Donnelly [3, Corollary 3.4]). It then follows that a major outbreak occurs with non-zero probability if and only if $R_0 > 1$, and the probability of a major outbreak is $1 - z^m$. However, for practical applications it is useful to have a definition of a major outbreak that depends on n. We say that a major outbreak occurs if and only if $T^{(n)} \ge \log n$.

Theorem 3.2. *Suppose that (3.9) holds and (3.10) holds for some $\theta_0 > 0$.*

(a) *Then*

$$\mathbb{P}\left(T^{(n)} \ge \log n\right) \to 1 - z^m \quad \text{as } n \to \infty. \tag{3.13}$$

(b) *If also $R_0 > 1$ and there exists $\alpha > 1$ such that $\mathbb{E}[C^{1+\alpha}] < \infty$, then there exists $\delta > 0$ such that*

$$\mathbb{P}\left(T^{(n)} \ge \delta n | T^{(n)} \ge \log n\right) \to 1 \quad \text{as } n \to \infty. \tag{3.14}$$

The proof of Theorem 3.2 is presented in Section 5.

Theorem 3.2(b) implies that a major outbreak infects at least a fraction δ of the population with probability tending to one as $n \to \infty$. However, δ depends on the parameters of the epidemic $\mathcal{E}^{(n)}$ and can be arbitrarily close to 0. An immediate consequence of the proof of Theorem 3.2(b) is Corollary 3.2, which states that, in the limit as $n \to \infty$, there exists $\delta' > 0$ such that in the event of a major epidemic outbreak the proportion of the population infectious at the peak of the epidemic exceeds δ'.

Corollary 3.2. *Under the conditions of Theorem 3.2(b), there exists $\delta' > 0$ such that*

$$\mathbb{P}\left(\sup_{t \geq 0} \left|\frac{I^{(n)}(t)}{n}\right| \geq \delta' \,\middle|\, T^{(n)} \geq \log n\right) \to 1 \quad \text{as } n \to \infty. \tag{3.15}$$

A central limit theorem for the total size $T^{(n)}$ in the event of a major outbreak is given in Ball and Neal [5], for the special case of binomial sampling (Section 2.1.2), by using the theory of (asymptotically) density-dependent population processes (Ethier and Kurtz [10, Chapter 11] and Pollett [16]) to obtain a functional central limit theorem for a random time-scale transformation of $\{(S^{(n)}(t), I^{(n)}(t)) : t \geq 0\}$ and hence a central limit theorem for the number of susceptibles when the number of infectives reaches zero, via a boundary-crossing problem. As noted in the introduction, Theorem 3.2(b) is a key step in the proof of the above central limit theorem, though the result was only stated in [5]. A similar central limit theorem for $T^{(n)}$ is likely to hold for our more general model, although details will be messy unless $\pi_c(w; i, j)$ takes a convenient form.

4. Proof of Theorem 3.2

4.1. Overview

We present an overview of the steps to prove Theorem 3.1. We construct on a common probability space the Markovian branching process $\mathcal{B}$ and the sequence of epidemic processes $(\mathcal{E}^{(n)})$, in which we equate infection and removal events in the epidemic process, $\mathcal{E}^{(n)}$, with birth and death events, respectively, in the branching process, $\mathcal{B}$. Given that at time $t \geq 0$ there are the same number of infectious individuals in the epidemic process $\mathcal{E}^{(n)}$ as there are individuals alive in the branching process $\mathcal{B}$, the removal rate in $\mathcal{E}^{(n)}$ is equal to the death rate in $\mathcal{B}$. For $k = 0, 1, \ldots$, the rate at which an infection event occurs which generates k new infections in $\mathcal{E}^{(n)}$ will depend upon the state of the population (number of susceptibles and infectives), and during the early stages of the epidemic this rate will be close to, but typically not equal to, the rate at which a birth event resulting in k new individuals occurs in $\mathcal{B}$. Therefore, we look to bound the difference between the infection rate in $\mathcal{E}^{(n)}$ and the birth rate in $\mathcal{B}$ in order to establish a coupling between the two processes.

A useful observation is that in the epidemic processes $\mathcal{E}^{(n)}$ (the branching process $\mathcal{B}$) the number of infectives and susceptibles (the number of individuals alive) is piecewise constant between events, where an event is either a mixing event or a recovery. Therefore, in Section 4.2, we define embedded discrete-time jump processes for $\mathcal{E}^{(n)}$ and $\mathcal{B}$, for the number of infectives (and susceptibles) and the number of individuals alive after each event. In the case of $\mathcal{B}$ the embedded discrete-time jump process is a random walk. Then, in Section 4.3, we provide a bound on the rate of convergence to 0 of the difference between the infection rate in $\mathcal{E}^{(n)}$ and the birth rate in $\mathcal{B}$ in Lemma 4.1, which is applicable during the early stages of the epidemic when only a few individuals have been infected. Lemma 4.1 enables us to construct the embedded discrete-time jump processes defined in Section 4.2 on a common probability space

(Section 4.4) and provide an almost sure coupling between the discrete-time processes during the initial stages of the epidemic (Section 4.5). That is, we couple the outcomes of the kth ($k = 1, 2, \ldots$) events in $\mathcal{E}^{(n)}$ and $\mathcal{B}$ so that the types of event—birth (infection) and death (removal)—match, and in the case of birth/infection, the same numbers of births and infections occur. Once we have established an almost sure agreement between the types of events that have occurred in the epidemic and branching processes, it is straightforward to provide an almost sure coupling of the timing of the events. The key couplings are drawn together in Lemma 4.2, from which Theorem 3.1 follows almost immediately. Finally, we consider the total sizes of the epidemic processes $\mathcal{E}^{(n)}$ and the branching process $\mathcal{B}$ and provide a proof of Corollary 3.1.

4.2. Embedded random walk

Let the random walk $\mathcal{R}$ be defined as follows. Let Y_k denote the position of the random walk after k steps, with $Y_0 = m > 0$. For $k = 1, 2, \ldots$, let $Y_k = Y_{k-1} + Z_k$, where $Z_1, Z_2, \ldots$ are i.i.d. with probability mass function

$$\mathbb{P}\left(Z_k = w\right) = \begin{cases} \dfrac{\beta \varphi_w}{\gamma + \beta}, & w = 0, 1, \ldots, \\ \dfrac{\gamma}{\gamma + \beta}, & w = -1, \\ 0 & \text{otherwise,} \end{cases} \tag{4.1}$$

where $\beta = \lambda \mu_C$ and φ_w is defined as in (3.3). Thus, upward (downward) steps in $\mathcal{R}$ correspond to birth (death) events in $\mathcal{B}$. Note that $Z_k = 0$ is possible, corresponding to a step with no change in the random walk (a birth event with no births in $\mathcal{B}$). For $k = 1, 2, \ldots$, let η_k denote the time of the kth event in $\mathcal{B}$ with $\eta_0 = 0$; then we can construct $\mathcal{R}$ from $\mathcal{B}$ by setting $Y_k = I(\eta_k)$, where $I(t)$ ($t \geq 0$) is the size of the population of $\mathcal{B}$ at time t. Note that if $I(\eta_k) = 0$, then the branching process has gone extinct and $Y_k = 0$, i.e. the random walk has hit 0. We can continue the construction of the random walk after the branching process has gone extinct using $Y_k = Y_{k-1} + Z_k$, but our primary interest is in the case where the two processes are positive. Conversely, we can construct $\mathcal{B}$ from $\mathcal{R}$ by using, in addition to $\{Y_k\} = \{Y_k : k = 0, 1, \ldots\}$, a sequence of i.i.d. random variables $V_1, V_2, \ldots$, where $V_k \sim \text{Exp}(1)$. (Throughout the paper, discrete-time processes are assumed to have index set $\mathbb{Z}_+$ unless indicated otherwise.) For $k = 1, 2, \ldots$,

$$\eta_k = \eta_{k-1} + \frac{V_k}{(\gamma + \beta) Y_{k-1}},$$

and for any $\eta_k \leq t < \eta_{k+1}$, we set $I(t) = Y_k$. Note that $\eta_k = \infty$ if $Y_{k-1} = 0$, corresponding to the branching process going extinct with $I(t) = 0$ for all $t \geq \eta_{k-1}$. Finally, note that $\mathbb{E}[Z_1] < 0$, $= 0$, or > 0 if and only if $R_0 < 1$, $= 1$, or > 1, respectively.

We turn to the sequence of epidemic processes, $(\mathcal{E}^{(n)})$, and for each $\mathcal{E}^{(n)}$, an associated discrete-time epidemic jump process $\mathcal{S}^{(n)}$. Let $Q_c^{(n)}(i, j | x, y)$ denote the probability that a mixing event of size c in a population of size n with x susceptibles and y infectives (and hence $n - x - y$ recovered individuals) involves i susceptibles and j infectives (and hence $c - i - j$ recovered individuals). Note that

$$Q_c^{(n)}(i, j | x, y) = \frac{\binom{x}{i} \times \binom{y}{j} \times \binom{n-x-y}{c-i-j}}{\binom{n}{c}}. \tag{4.2}$$

For $w = 0, 1, \ldots$, let $q^{(n)}(x, y, w)$ be such that

$$q^{(n)}(x, y, w)y = n\lambda \sum_{c=w+1}^{n} p_C^{(n)}(c) \sum_{j=1}^{c-w} \sum_{l=0}^{c-w-j} \left\{ Q_c^{(n)}(c - j - l, j | x, y)\pi_c(w; c - j - l, j) \right\},$$

(4.3)

where the indices j and l refer to the numbers of infectives and recovered individuals, respectively, involved in the mixing event. Thus, for $w = 1, 2, \ldots, x$, $q^{(n)}(x, y, w)y$ denotes the rate of occurrence of mixing events that create w new infectives within a population of size n having x susceptibles and y infectives. Hence, $q^{(n)}(x, y, w)$ can be viewed as the rate at which an infectious individual in a population of size n containing x susceptibles and y infectives generates w new infectives. Note that $q^{(n)}(x, y, 0)y$ is the rate of occurrence of mixing events which involve at least one infective and create no new infectives, in a population with x susceptibles and $y > 0$ infectives.

Recall that, for $t \geq 0$, $S^{(n)}(t)$ and $I^{(n)}(t)$ denote respectively the numbers of susceptibles and infectives at time t in $\mathcal{E}^{(n)}$. Since the population is closed, for all $t \geq 0$, $n - S^{(n)}(t) - I^{(n)}(t)$ denotes the number of recovered individuals, and we can describe the epidemic $\mathcal{E}^{(n)}$ in terms of $\left\{ (S^{(n)}(t), I^{(n)}(t)) : t \geq 0 \right\}$, which is a continuous-time Markov chain on the state space $E^{(n)} = \left\{ (x, y) \in \mathbb{Z}^2 : x + y \leq n, 0 \leq x \leq n - m_n, y \geq 0 \right\}$ with transition probabilities

$$\mathbb{P}\left((S^{(n)}(t + \Delta t), I^{(n)}(t + \Delta t)) = (x - w, y + w) \middle| (S^{(n)}(t), I^{(n)}(t)) = (x, y) \right)$$
$$= q^{(n)}(x, y, w)y\Delta t + o(\Delta t) \qquad (w = 0, 1, \ldots, x), \qquad (4.4)$$

$$\mathbb{P}\left((S^{(n)}(t + \Delta t), I^{(n)}(t + \Delta t)) = (x, y - 1) \middle| (S^{(n)}(t), I^{(n)}(t)) = (x, y) \right)$$
$$= \gamma y\Delta t + o(\Delta t), \qquad (4.5)$$

and with all other transitions having probability $o(\Delta t)$. The events (4.4) and (4.5) correspond to infection of w individuals and recovery of an individual, respectively. The function $q^{(n)}$ is real-valued with domain a subset of $\mathbb{Z}_+ \times \mathbb{Z}_+ \times \mathbb{N}$. We note that the epidemic process is invariant to the choice of $q^{(n)}(x, y, 0) \geq 0$, so we can define $q^{(n)}(x, y, 0)$ to satisfy (4.3) with $w = 0$. Similarly, the epidemic process is invariant to the choice of $q^{(n)}(x, 0, w)$, as no infections can occur if $y = 0$, but for coupling purposes it is useful to define $q^{(n)}(x, y, w) = \beta\varphi_w$ for $y = 0, -1, -2, \ldots$. Finally, as noted in Section 4.1, we observe that the recovery rate (4.5) coincides with the death rate of the branching process $\mathcal{B}$, so to couple the number of infectives in the epidemic process $\mathcal{E}^{(n)}$ to the number of individuals in the branching process $\mathcal{B}$, we require that $q^{(n)}(x, y, w) \approx \beta\varphi_w$ and $q^{(n)}(x, y) = \sum_{w=0}^{n-1} q^{(n)}(x, y, w) = \sum_{w=0}^{\infty} q^{(n)}(x, y, w) \approx \beta$ as n becomes large. $\left(\text{Note that for } w > n - 1, q^{(n)}(x, y, w) = 0.\right)$ We proceed by making this precise after first describing the discrete-time epidemic jump process $\mathcal{S}^{(n)}$.

For $n = 1, 2, \ldots$ and $k = 0, 1, \ldots$, let $\left(X_k^{(n)}, Y_k^{(n)} \right)$ denote the state of the jump process $\mathcal{S}^{(n)}$ after the kth event with $\left(X_0^{(n)}, Y_0^{(n)} \right) = (n - m_n, m_n)$. For $k = 1, 2, \ldots, (x, y) \in E^{(n)}$, and $w = 0, 1, \ldots, x$, let

$$\mathbb{P}\left(\left(X_{k+1}^{(n)}, Y_{k+1}^{(n)} \right) = (x - w, y + w) \middle| \left(X_k^{(n)}, Y_k^{(n)} \right) = (x, y) \right) = \frac{q^{(n)}(x, y, w)}{\gamma + q^{(n)}(x, y)},$$

$$\mathbb{P}\left(\left(X_{k+1}^{(n)}, Y_{k+1}^{(n)} \right) = (x, y - 1) \middle| \left(X_k^{(n)}, Y_k^{(n)} \right) = (x, y) \right) = \frac{\gamma}{\gamma + q^{(n)}(x, y)},$$

with all other transitions having probability 0 of occurring. Letting $\eta_k^{(n)}$ denote the time of the kth event in $\mathcal{E}^{(n)}$, with $\eta_0^{(n)} = 0$, we can construct $\mathcal{S}^{(n)}$ from $\mathcal{E}^{(n)}$ by setting $\left(X_k^{(n)}, Y_k^{(n)}\right) = \left(S^{(n)}\left(\eta_k^{(n)}\right), I^{(n)}\left(\eta_k^{(n)}\right)\right)$. As with the construction of $\mathcal{R}$, we can continue the construction of $\mathcal{S}^{(n)}$ after the kth event with $Y_k^{(n)} = 0$, using $q^{(n)}(x, y, w) = \beta \varphi_w$ for $y = 0, -1, -2, \ldots$. Conversely, we can construct $\mathcal{E}^{(n)}$ from $\mathcal{S}^{(n)}$ by in addition using the sequence of i.i.d. random variables $V_1^{(n)}, V_2^{(n)}, \ldots$, where $V_i^{(n)} \sim \mathrm{Exp}(1)$. For $k = 1, 2, \ldots$, set

$$\eta_k^{(n)} = \eta_{k-1}^{(n)} + \frac{V_k^{(n)}}{\left[\gamma + q^{(n)}\left(X_{k-1}^{(n)}, Y_{k-1}^{(n)}\right)\right] Y_{k-1}^{(n)}}, \tag{4.6}$$

then for any $\eta_k^{(n)} \leq t < \eta_{k+1}^{(n)}$, set $\left(S^{(n)}(t), I^{(n)}(t)\right) = \left(X_k^{(n)}, Y_k^{(n)}\right)$. Note that if $Y_{k-1}^n = 0$, $\eta_k^{(n)} = \infty$ and for all $t \geq \eta_{k-1}^{(n)}$, the epidemic has died out with $\left(S^{(n)}(t), I^{(n)}(t)\right) = \left(X_{k-1}^{(n)}, 0\right)$.

We briefly discuss the choice of $V_k^{(n)}$. A simple coupling with the branching process $\mathcal{B}$ would be to set $V_k^{(n)} = V_k$, which results in $\eta_k^{(n)} \approx \eta_k$ if $\eta_{k-1}^{(n)} \approx \eta_{k-1}$ and $Y_{k-1}^{(n)}\left[\gamma + q^{(n)}\left(X_{k-1}^{(n)}, Y_{k-1}^{(n)}\right)\right] \approx Y_{k-1}[\gamma + \beta]$. This is the approach taken in [6] and leads to a slight mismatch between the event times in $\mathcal{E}^{(n)}$ and $\mathcal{B}$, with the mismatch converging to 0 as $n \to \infty$. Therefore we take an alternative approach which results in there being high probability of $\eta_k^{(n)} = \eta_k$, if $\eta_{k-1}^{(n)} = \eta_{k-1}$ and $Y_{k-1}^{(n)} = Y_{k-1}$, with the details provided in Section 4.4.

4.3. Matching infection rate to birth rate

In this section, we provide bounds on the differences between the rate, $q^{(n)}\left(x^{(n)}, y^{(n)}, w\right)$, at which events creating w ($w = 0, 1, \ldots$) new infectives occur in $\mathcal{E}^{(n)}$ with $x^{(n)}$ susceptibles and $y^{(n)}$ infectives present in the population, and the rate, $\beta \varphi_w$, at which birth events creating w new individuals occur in $\mathcal{B}$. The bounds on the difference in rates are appropriate during the early stages of the epidemic process where $n - r_n \leq x \leq n - m_n$ (i.e. whilst fewer than r_n individuals have ever been in the infectious state), for a sequence (r_n) satisfying $r_n \to \infty$ and $r_n / \sqrt{n} \to 0$ as $n \to \infty$.

In the early stages of the epidemic, when $x \geq n - r_n$, it is unlikely that a mixing event will involve more than one non-susceptible individual. Thus we split the double sum over j and l in (4.3) into the case $j = 1$ and $l = 0$, a single infective in an otherwise susceptible group of size c, and the case $j + l \geq 2$, where there is more than one non-susceptible individual in a mixing event. This gives, for $y > 0$,

$$q^{(n)}(x, y, w) = \frac{n\lambda}{y} \sum_{c=w+1}^{n} p_C^{(n)}(c) Q_c^{(n)}(c - 1, 1 | x, y) \pi_c(w; c - 1, 1)$$

$$+ \frac{n\lambda}{y} \sum_{c=w+1}^{n} p_C^{(n)}(c) \sum_{j+l \geq 2} \left\{ Q_c^{(n)}(c - j - l, j | x, y) \pi_c(w; c - j - l, j) \right\}$$

$$= q_1^{(n)}(x, y, w) + q_2^{(n)}(x, y, w), \quad \text{say.} \tag{4.7}$$

We consider the two terms on the right-hand side of (4.7). Note that for $y \leq 0$, we set $q_1^{(n)}(x, y, w) = \beta \varphi_w$ and $q_2^{(n)}(x, y, w) = 0$, which is consistent with $q^{(n)}(x, y, w) = \beta \varphi_w$

$(y = 0, -1, \dots)$. Also, for $w = n, n+1, \dots,$ $q^{(n)}(x, y, w) = 0$, which implies $q_h^{(n)}(x, y, w) = 0$ $(h = 1, 2)$. For $h = 1, 2$, let $q_h^{(n)}(x, y) = \sum_{w=0}^{n-1} q_h^{(n)}(x, y, w) = \sum_{w=0}^{\infty} q_h^{(n)}(x, y, w)$, the sums over w of the two components of $q^{(n)}(x, y, w)$ in (4.7). Hence $q^{(n)}(x, y) = q_1^{(n)}(x, y) + q_2^{(n)}(x, y)$.

Lemma 4.1 provides bounds on the rate of convergence to 0, as $n \to \infty$, of the difference between the infection rate in the epidemic process and the birth rate in the branching process, in terms of the number of non-susceptibles in the population (r_n) and the rate of convergence of $C^{(n)}$ to C. Remember that throughout we assume that $C^{(n)} \xrightarrow{\text{D}} C$ and $\mathbb{E}[(C^{(n)})^2] \to \mathbb{E}[C^2]$ as $n \to \infty$, with $\mathbb{E}[C^2] < \infty$; see the conditions stated before Theorem 3.1 in Section 3.2.

Lemma 4.1. *Let (r_n) be a sequence of positive real numbers such that $r_n \to \infty$ and $r_n/\sqrt{n} \to 0$ as $n \to \infty$.*

Let (s_n) be a sequence of positive real numbers such that $s_n r_n^2/n \to 0$ and

$$s_n \sum_{c=2}^{n} c \left| p_C^{(n)}(c) - p_C(c) \right| \to 0 \quad \text{as } n \to \infty. \tag{4.8}$$

Suppose that $(x^{(n)})$ and $(y^{(n)})$ are two sequences such that $n - r_n \le x^{(n)} \le n - m_n$ and $0 < y^{(n)} \le r_n$ for all sufficiently large n. Then

$$s_n \sum_{w=0}^{\infty} \left| q_1^{(n)}(x^{(n)}, y^{(n)}, w) - \beta \varphi_w \right| \to 0 \quad \text{as } n \to \infty \tag{4.9}$$

and

$$\sum_{w=0}^{\infty} s_n \left| q_2^{(n)}(x^{(n)}, y^{(n)}, w) \right| = s_n \sum_{w=0}^{n-1} \left| q_2^{(n)}(x^{(n)}, y^{(n)}, w) \right| \to 0 \quad \text{as } n \to \infty. \tag{4.10}$$

Consequently, if $s_n r_n^2/n \to 0$ as $n \to \infty$, then

$$s_n \left| q^{(n)}(x^{(n)}, y^{(n)}) - \beta \right| = s_n \left| \sum_{w=0}^{\infty} \left\{ q^{(n)}(x^{(n)}, y^{(n)}, w) - \beta \varphi_w \right\} \right| \to 0 \quad \text{as } n \to \infty. \tag{4.11}$$

Proof. First note that, for $Q_c^{(n)}(c - 1, 1 | x^{(n)}, y^{(n)})$ defined in (4.2) and any $c = 2, 3, \dots,$

$$Q_c^{(n)}(c - 1, 1 | x^{(n)}, y^{(n)}) = c \frac{y^{(n)}}{n} \prod_{j=0}^{c-2} \frac{x^{(n)} - j}{n - 1 - j}$$

$$= c \frac{y^{(n)}}{n} + y^{(n)} \epsilon_c^{(n)}(x^{(n)}), \tag{4.12}$$

where

$$\epsilon_c^{(n)}(x^{(n)}) = \frac{c}{n} \left\{ \left[\prod_{j=0}^{c-2} \frac{x^{(n)} - j}{n - 1 - j} \right] - 1 \right\}.$$

For $x^{(n)} \geq n - r_n$ and $c \leq n/2$, we have that

$$
1 \geq \prod_{j=0}^{c-2} \frac{x^{(n)} - j}{n - 1 - j} = \prod_{j=0}^{c-2} \left(1 - \frac{n - 1 - x^{(n)}}{n - 1 - j} \right)
$$

$$
\geq 1 - [n - 1 - x^{(n)}] \left\{ \frac{1}{n - 1} + \frac{1}{n - 2} + \ldots + \frac{1}{n - c + 1} \right\}
$$

$$
\geq 1 - [r_n - 1] \left\{ \frac{1}{n - 1} + \frac{1}{n - 2} + \ldots + \frac{1}{n - c + 1} \right\}
$$

$$
\geq 1 - [r_n - 1] \frac{c - 1}{n - c + 1} \geq 1 - \frac{2 r_n (c - 1)}{n}.
$$

Therefore, for $x^{(n)} \geq n - r_n$ and $c \leq n/2$,

$$
-\frac{c}{n} \times \frac{2 r_n (c - 1)}{n} \leq \epsilon_c^{(n)}(x^{(n)}) \leq 0. \tag{4.13}
$$

Note that $p_C^{(n)}(c) = 0$ for $c > n$. Also, using (3.3) and recalling that $\beta = \lambda \mu_C$, we have

$$
\lambda \sum_{c=w+1}^{\infty} c p_C(c) \pi_c(w; c - 1, 1) = \lambda \mu_C \varphi_w = \beta \varphi_w \qquad (w = 0, 1, \ldots).
$$

Hence, for $w = 0, 1, \ldots,$

$$
q_1^{(n)}\left(x^{(n)}, y^{(n)}, w \right) = \frac{n\lambda}{y^{(n)}} \sum_{c=w+1}^{n} p_C^{(n)}(c) Q_c^{(n)}\left(c - 1, 1 | x^{(n)}, y^{(n)} \right) \pi_c(w; c - 1, 1)
$$

$$
= \frac{n\lambda}{y^{(n)}} \sum_{c=w+1}^{n} p_C^{(n)}(c) \left\{ \frac{c y^{(n)}}{n} + y^{(n)} \epsilon_c^{(n)}(x^{(n)}) \right\} \pi_c(w; c - 1, 1)
$$

$$
= \lambda \sum_{c=w+1}^{n} c p_C^{(n)}(c) \pi_c(w; c - 1, 1)
$$

$$
+ \lambda \sum_{c=w+1}^{n} p_C^{(n)}(c) n \epsilon_c^{(n)}(x^{(n)}) \pi_c(w; c - 1, 1)
$$

$$
= \lambda \sum_{c=w+1}^{\infty} c p_C(c) \pi_c(w; c - 1, 1)
$$

$$
+ \lambda \sum_{c=w+1}^{\infty} c \left[p_C^{(n)}(c) - p_C(c) \right] \pi_c(w; c - 1, 1)
$$

$$
+ \lambda \sum_{c=w+1}^{n} p_C^{(n)}(c) n \epsilon_c^{(n)}(x^{(n)}) \pi_c(w; c - 1, 1)
$$

$$
= \beta \varphi_w + \lambda \sum_{c=w+1}^{\infty} c \left[p_C^{(n)}(c) - p_C(c) \right] \pi_c(w; c - 1, 1)
$$

$$
+ \lambda \sum_{c=w+1}^{n} p_C^{(n)}(c) n \epsilon_c^{(n)}(x^{(n)}) \pi_c(w; c - 1, 1).
$$

It follows that

$$\sum_{w=0}^{\infty} s_n \left| q_1^{(n)}(x^{(n)}, y^{(n)}, w) - \beta\varphi_w \right|$$

$$\leq s_n \sum_{w=0}^{\infty} \left| \lambda \sum_{c=w+1}^{\infty} c[p_C^{(n)}(c) - p_C(c)]\pi_c(w; c-1, 1) \right|$$

$$+ s_n \sum_{w=0}^{\infty} \left| \lambda \sum_{c=w+1}^{n} p_C^{(n)}(c) n\epsilon_c^{(n)}(x^{(n)})\pi_c(w; c-1, 1) \right|$$

$$\leq \lambda s_n \sum_{w=0}^{\infty} \sum_{c=w+1}^{\infty} c\left|p_C^{(n)}(c) - p_C(c)\right|\pi_c(w; c-1, 1)$$

$$+ \lambda s_n \sum_{w=0}^{\infty} \sum_{c=w+1}^{n} p_C^{(n)}(c) \left|n\epsilon_c^{(n)}(x^{(n)})\right| \pi_c(w; c-1, 1)$$

$$= \lambda s_n \sum_{c=2}^{\infty} c\left|p_C^{(n)}(c) - p_C(c)\right| \sum_{w=0}^{c-1} \pi_c(w; c-1, 1)$$

$$+ \lambda s_n \sum_{c=2}^{n} p_C^{(n)}(c) \left|n\epsilon_c^{(n)}(x^{(n)})\right| \sum_{w=0}^{c-1} \pi_c(w; c-1, 1)$$

$$= \lambda s_n \sum_{c=2}^{\infty} c\left|p_C^{(n)}(c) - p_C(c)\right| + \lambda s_n \sum_{c=2}^{n} p_C^{(n)}(c) \left|n\epsilon_c^{(n)}(x^{(n)})\right|. \qquad (4.14)$$

The first term on the right-hand side of (4.14) converges to 0 by (4.8). Using (4.13) and Markov's inequality, the second term on the right-hand side of (4.14) satisfies

$$\lambda s_n \sum_{c=2}^{n} p_C^{(n)}(c) \left|n\epsilon_c^{(n)}(x^{(n)})\right|$$

$$\leq \lambda \left\{ \sum_{c=2}^{\lfloor n/2 \rfloor} s_n p_C^{(n)}(c) \left(\frac{2c(c-1)r_n}{n}\right) + \sum_{c=\lfloor n/2 \rfloor+1}^{n} s_n p_C^{(n)}(c)c \right\}$$

$$\leq \lambda \left\{ 2\frac{s_n r_n}{n}\mathbb{E}\left[C^{(n)}(C^{(n)} - 1)\right] + s_n n\mathbb{P}\left(C^{(n)} > \lfloor n/2 \rfloor + 1\right) \right\}$$

$$\leq \lambda \left\{ 2\frac{s_n r_n}{n}\mathbb{E}\left[C^{(n)}(C^{(n)} - 1)\right] + s_n n \times \frac{4}{n^2}\mathbb{E}\left[\left(C^{(n)}\right)^2\right] \right\}$$

$$\to 0 \quad \text{as } n \to \infty.$$

Hence (4.9) is proved.

The probability that a pair of individuals, chosen uniformly at random, are both non-susceptible is $(n - x^{(n)})(n - x^{(n)} - 1)/[n(n-1)]$. In a group of c individuals there are $c(c-1)/2$ pairs, so

$$\sum_{j+l\geq 2} Q_c^{(n)}(c - j - l, j|x^{(n)}, y^{(n)}) \leq \frac{c(c-1)}{2} \times \frac{(n - x^{(n)})(n - x^{(n)} - 1)}{n(n-1)}. \qquad (4.15)$$

For $x^{(n)} \geq n - r_n$, the right-hand side of (4.15) is bounded above by $[c(c-1)/2] \times [r_n/n]^2$.
Therefore, since $q_2^{(n)}(x^{(n)}, y^{(n)}, w) = 0$ for $w = n, n+1, \ldots$, we have that

$$
\sum_{w=0}^{\infty} s_n \left| q_2^{(n)}(x^{(n)}, y^{(n)}, w) \right| = \sum_{w=0}^{n-1} s_n \left| q_2^{(n)}(x^{(n)}, y^{(n)}, w) \right|
$$

$$
= \frac{s_n n}{y^{(n)}} \lambda \sum_{w=0}^{n-1} \sum_{c=w+1}^{n} p_C^{(n)}(c) \sum_{j+l \geq 2} \left\{ Q_c^{(n)}(c-j-l, j | x^{(n)}, y^{(n)}) \pi_c(w; c-j-l, j) \right\}
$$

$$
= \frac{s_n n}{y^{(n)}} \lambda \sum_{c=2}^{n} p_C^{(n)}(c) \sum_{j+l \geq 2} \left\{ Q_c^{(n)}(c-j-l, j | x^{(n)}, y^{(n)}) \sum_{w=0}^{c-j-l} \pi_c(w; c-j-l, l, j) \right\}
$$

$$
\leq \lambda \frac{s_n r_n^2}{n} \sum_{c=2}^{n} p_C^{(n)}(c) \frac{c(c-1)}{2}
$$

$$
= \lambda \frac{s_n r_n^2}{2n} \mathbb{E}\left[C^{(n)}(C^{(n)} - 1) \right] \to 0 \quad \text{as } n \to \infty,
$$

and (4.10) is proved.

Finally, (4.11) follows from (4.9) and (4.10) by the triangle inequality. $\qquad \square$

Note that if C has finite support $\{2, 3, \ldots, n_0\}$, then for all $n \geq n_0$, $C^{(n)} \equiv C$, and (4.8) holds for any sequence $\{s_n\}$.

4.4. Construction of the event processes

Lemma 4.1 implies that the difference between the transition probabilities of $\mathcal{R}$ and $\mathcal{S}^{(n)}$ tends to 0 as $n \to \infty$, provided the number of non-susceptible individuals remains sufficiently small. We proceed by constructing $\mathcal{R}$ and $\mathcal{S}^{(n)}$ on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with $Y_0 = m$ and, for all sufficiently large n, $\left(X_0^{(n)}, Y_0^{(n)} \right) = (n - m_n, m_n) = (n - m, m)$. For $k = 1, 2, \ldots$, let $\mathbf{U}_k = (U_{k,1}, U_{k,2}, U_{k,3})$ be i.i.d. random vectors defined on $(\Omega, \mathcal{F}, \mathbb{P})$, with $U_{k,i} \sim U(0, 1)$ ($i = 1, 2, 3$) being independent, where $U(0, 1)$ denotes a random variable that is uniformly distributed on $[0, 1]$.

We construct $\mathcal{R}$ as follows. Suppose that for some $k = 1, 2, \ldots$, $Y_{k-1} = y$. The kth step in $\mathcal{R}$ is a downward step (of size 1) with $Y_k = y - 1$ if $U_{k,1} \leq \gamma/(\gamma + \beta)$. Otherwise the random walk has an 'upward' step of size a_k with $Y_k = y + a_k$, where a_k satisfies

$$
\sum_{l=0}^{a_k - 1} \varphi_l < U_{k,2} \leq \sum_{l=0}^{a_k} \varphi_l.
$$

Note that all sums are equal to 0, if vacuous; $a_k = 0$ is possible and the probability that $a_k = i$ is φ_i.

Similarly, we construct $\mathcal{S}^{(n)}$ as follows. Suppose that for some $k = 1, 2, \ldots$, $\left(X_{k-1}^{(n)}, Y_{k-1}^{(n)} \right) = \left(x_k^{(n)}, y_k^{(n)} \right)$. The kth event in $\mathcal{S}^{(n)}$ is a recovery with $\left(X_k^{(n)}, Y_k^{(n)} \right) = \left(x_k^{(n)}, y_k^{(n)} - 1 \right)$ if $U_{k,1} \leq \gamma/\left[\gamma + q^{(n)}\left(x_k^{(n)}, y_k^{(n)} \right) \right]$. Otherwise the kth event in $\mathcal{S}^{(n)}$ is an infection event of size $a_k^{(n)}$ with $\left(X_k^{(n)}, Y_k^{(n)} \right) = \left(x_k^{(n)} - a_k^{(n)}, y_k^{(n)} + a_k^{(n)} \right)$, where $a_k^{(n)}$ satisfies

$$\mathbb{P}\big(a_k^{(n)} = i\big) = \frac{q^{(n)}\big(x_k^{(n)}, y_k^{(n)}, i\big)}{q^{(n)}\big(x_k^{(n)}, y_k^{(n)}\big)} = \varphi_i^{(n)}\big(x_k^{(n)}, y_k^{(n)}\big), \quad \text{say.}$$

To enable an effective coupling between $\mathcal{R}$ and $\mathcal{S}^{(n)}$, we obtain $a_k^{(n)}$ as follows. For $i = 0, 1, \ldots$, let $\varpi_i^{(n)}\big(x_k^{(n)}, y_k^{(n)}\big) = \min\left\{\varphi_i, \varphi_i^{(n)}\big(x_k^{(n)}, y_k^{(n)}\big)\right\}$ and let

$$D_2^{(n)}\big(x_k^{(n)}, y_k^{(n)}\big) = \bigcup_{w=0}^{\infty} \left(\sum_{l=0}^{w-1} \varphi_l + \varpi_w^{(n)}\big(x_k^{(n)}, y_k^{(n)}\big), \sum_{l=0}^{w} \varphi_l\right], \tag{4.16}$$

where $(a, b]$ is the empty set if $a = b$. If $U_{k,2} \notin D_2^{(n)}\big(x_k^{(n)}, y_k^{(n)}\big)$, then there exists $i \in \mathbb{Z}_+$ such that

$$\sum_{l=0}^{i-1} \varphi_l < U_{k,2} \le \sum_{l=0}^{i-1} \varphi_l + \varpi_i^{(n)}\big(x_k^{(n)}, y_k^{(n)}\big), \tag{4.17}$$

and we set $a_k^{(n)} = i$. Therefore, if $U_{k,2} \notin D_2^{(n)}\big(x_k^{(n)}, y_k^{(n)}\big)$, we have that $a_k^{(n)} = a_k$. Let

$$d_k^{(n)}\big(x_k^{(n)}, y_k^{(n)}\big) = \mathbb{P}\left(U_{k,2} \in D_2^{(n)}\big(x_k^{(n)}, y_k^{(n)}\big)\right) = 1 - \sum_{v=0}^{\infty} \varpi_v^{(n)}\big(x_k^{(n)}, y_k^{(n)}\big),$$

the total variation distance between $(\varphi_0, \varphi_1, \ldots)$ and $\left(\varphi_0^{(n)}\big(x_k^{(n)}, y_k^{(n)}\big), \varphi_1^{(n)}\big(x_k^{(n)}, y_k^{(n)}\big), \ldots\right)$. If $U_{k,2} \in D_2^{(n)}\big(x_k^{(n)}, y_k^{(n)}\big)$, we set $a_k^{(n)} = i$ with probability

$$\frac{\varphi_i^{(n)}\big(x_k^{(n)}, y_k^{(n)}\big) - \varpi_i^{(n)}\big(x_k^{(n)}, y_k^{(n)}\big)}{d_k^{(n)}\big(x_k^{(n)}, y_k^{(n)}\big)},$$

which ensures that overall the probability that $a_k^{(n)} = i$ is $\varphi_i^{(n)}\big(x_k^{(n)}, y_k^{(n)}\big)$. We do not need to be more explicit about the choice $a_k^{(n)}$ when $a_k^{(n)} \ne a_k$.

Given $V_1, V_2, \ldots$, i.i.d. according to Exp(1), we can construct $\mathcal{B}$ from $\mathcal{R}$ as outlined in Section 4.2. We conclude this section with a description of the construction of $\mathcal{E}^{(n)}$ from $\mathcal{S}^{(n)}$, in order to couple the time of events in $\mathcal{E}^{(n)}$ to the event times in $\mathcal{B}$. Given that there are $y^{(n)}$ infectives in the population, the probability that an individual chosen uniformly at random is infectious is $y^{(n)}/n$, so the probability that a mixing event of size c involves at least one infective is bounded above by $cy^{(n)}/n$. Therefore

$$q^{(n)}\big(x^{(n)}, y^{(n)}\big) = \sum_{w=0}^{n-1} q^{(n)}\big(x^{(n)}, y^{(n)}, w\big) \le \frac{1}{y^{(n)}} n\lambda \sum_{c=2}^{n} \frac{cy^{(n)}}{n} \times p_C^{(n)}(c)$$
$$= \lambda \mu_C^{(n)}. \tag{4.18}$$

Hence, under the assumption $\mu_C^{(n)} \le \mu_C$, we have that $q^{(n)}\big(x^{(n)}, y^{(n)}\big) \le \beta(= \lambda \mu_C)$. Therefore, letting

$$\tilde{d}^{(n)}\big(x^{(n)}, y^{(n)}\big) = \frac{\beta - q^{(n)}\big(x^{(n)}, y^{(n)}\big)}{\gamma + \beta} \ge 0, \tag{4.19}$$

we have, if $\left(X_{k-1}^{(n)}, Y_{k-1}^{(n)}\right) = \left(x^{(n)}, y^{(n)}\right)$, that

$$\frac{\gamma + \beta}{\gamma + q^{(n)}\left(X_{k-1}^{(n)}, Y_{k-1}^{(n)}\right)} V_k \sim \text{Exp}\left(1 - \tilde{d}^{(n)}\left(x^{(n)}, y^{(n)}\right)\right) = \tilde{V}_k^{(n)}\left(x^{(n)}, y^{(n)}\right), \text{ say.}$$

For $z \geq 0$, let

$$\tilde{f}_V\left(z; x^{(n)}, y^{(n)}\right) = \left\{1 - \tilde{d}^{(n)}\left(x^{(n)}, y^{(n)}\right)\right\} \exp\left(-z\left\{1 - \tilde{d}^{(n)}\left(x^{(n)}, y^{(n)}\right)\right\}\right) \qquad (4.20)$$

denote the probability density function of $\tilde{V}_k^{(n)}\left(x^{(n)}, y^{(n)}\right)$. Similarly, let $f_V(z) = \exp\left(-z\right)$ $(z \geq 0)$ denote the probability density function of V_1. It follows from (4.20), for all $z \geq 0$, that

$$\begin{aligned}
\tilde{f}_V\left(z; x^{(n)}, y^{(n)}\right) &\geq \left\{1 - \tilde{d}^{(n)}\left(x^{(n)}, y^{(n)}\right)\right\} \exp\left(-z\right) \\
&= \left\{1 - \tilde{d}^{(n)}\left(x^{(n)}, y^{(n)}\right)\right\} f_V(z).
\end{aligned}$$

Therefore, we can construct a realisation of $\tilde{V}_k^{(n)}\left(x^{(n)}, y^{(n)}\right)$ by setting $\tilde{V}_k^{(n)}\left(x^{(n)}, y^{(n)}\right) = V_k$ if $U_{k,3} \leq 1 - \tilde{d}^{(n)}\left(x^{(n)}, y^{(n)}\right)$, and if $U_{k,3} > 1 - \tilde{d}^{(n)}\left(x^{(n)}, y^{(n)}\right)$, we draw $\tilde{V}_k^{(n)}\left(x^{(n)}, y^{(n)}\right)$ from a random variable with, for $z \geq 0$, probability density function

$$f^*\left(z; x^{(n)}, y^{(n)}\right) = \frac{1 - \tilde{d}^{(n)}\left(x^{(n)}, y^{(n)}\right)}{\tilde{d}^{(n)}\left(x^{(n)}, y^{(n)}\right)} \left[\exp\left(-z\left\{1 - \tilde{d}^{(n)}\left(x^{(n)}, y^{(n)}\right)\right\}\right) - \exp\left(-z\right)\right]. \qquad (4.21)$$

Finally, we set

$$V_k^{(n)} = \tilde{V}_k^{(n)}\left(x^{(n)}, y^{(n)}\right) \frac{\gamma + q^{(n)}\left(x^{(n)}, y^{(n)}\right)}{\gamma + \beta}, \qquad (4.22)$$

which ensures that $V_k^{(n)} \sim \text{Exp}(1)$. Also, if $\eta_{k-1}^{(n)} = \eta_{k-1}$, $Y_{k-1}^{(n)} = Y_{k-1}$, and $U_{k,3} \leq 1 - \tilde{d}^{(n)}\left(X_{k-1}^{(n)}, Y_{k-1}^{(n)}\right)$, then $\tilde{V}_k^{(n)}\left(X_{k-1}^{(n)}, Y_{k-1}^{(n)}\right) = V_k$, and substituting $V_k^{(n)}$ into (4.6) and using (4.22) gives

$$\begin{aligned}
\eta_k^{(n)} &= \eta_{k-1}^{(n)} + \frac{V_k^{(n)}}{Y_{k-1}^{(n)}\left[\gamma + q^{(n)}\left(X_{k-1}^{(n)}, Y_{k-1}^{(n)}\right)\right]} \\
&= \eta_{k-1} + \frac{1}{Y_{k-1}\left[\gamma + q^{(n)}\left(X_{k-1}^{(n)}, Y_{k-1}^{(n)}\right)\right]} \times \frac{\gamma + q^{(n)}\left(X_{k-1}^{(n)}, Y_{k-1}^{(n)}\right)}{\gamma + \beta} \tilde{V}_k^{(n)}\left(X_{k-1}^{(n)}, Y_{k-1}^{(n)}\right) \\
&= \eta_{k-1} + \frac{V_k}{Y_{k-1}(\gamma + \beta)} = \eta_k.
\end{aligned} \qquad (4.23)$$

4.5. Coupling of the epidemic and branching processes

A *mismatch* occurs at event k whenever the kth events in the epidemic process, $\mathcal{E}^{(n)}$ (discrete epidemic jump process $\mathcal{S}^{(n)}$), and branching process, $\mathcal{B}$ (random walk $\mathcal{R}$), are either a removal ($\mathcal{E}^{(n)}$) and a birth ($\mathcal{B}$), or an infection ($\mathcal{E}^{(n)}$) and a birth ($\mathcal{B}$) where the number of new infections ($\mathcal{E}^{(n)}$) and the number of births ($\mathcal{B}$) differ. The first type of mismatch occurs in Ball and O'Neill [6], where also mismatches of the type with an infection ($\mathcal{E}^{(n)}$) and a death ($\mathcal{B}$) are permissible. Owing to (4.18) and the assumption that $\mu_C^{(n)} \leq \mu_C$ for all sufficiently large n, an infection in ($\mathcal{E}^{(n)}$) and a death ($\mathcal{B}$) is not possible for such n in the current setup, but the arguments can easily be modified to allow for this situation. The second type of mismatch comes from allowing multiple infections/births.

Since $q^{(n)}(x^{(n)}, y^{(n)}) \leq \beta$, a type-1 mismatch occurs at event k, where after event $k-1$ there are $x^{(n)}$ susceptibles and $y^{(n)}$ infectives, if and only if

$$U_{k,1} \in D_1^{(n)}(x^{(n)}, y^{(n)}) \equiv \left(\frac{\gamma}{\beta + \gamma}, \frac{\gamma}{q^{(n)}(x^{(n)}, y^{(n)}) + \gamma} \right], \tag{4.24}$$

with

$$\mathbb{P}\left(U_{k,1} \in D_1^{(n)}(x^{(n)}, y^{(n)}) \right) = \frac{\gamma \left[\beta - q^{(n)}(x^{(n)}, y^{(n)}) \right]}{\left[\gamma + q^{(n)}(x^{(n)}, y^{(n)}) \right] [\gamma + \beta]}. \tag{4.25}$$

Let $\tilde{Z}_1, \tilde{Z}_2, \ldots$ be i.i.d. according to $\tilde{Z}$ with probability mass function $\mathbb{P}(\tilde{Z} = i) = \varphi_i$ ($i = 0, 1, \ldots$). We construct $\tilde{Z}_1, \tilde{Z}_2, \ldots$ from $U_{1,2}, U_{2,2}, \ldots$ by setting $\tilde{Z}_k$ to satisfy

$$\sum_{i=0}^{\tilde{Z}_k - 1} \varphi_i < U_{k,2} \leq \sum_{i=0}^{\tilde{Z}_k} \varphi_i. \tag{4.26}$$

Thus $\tilde{Z}_k$ is the number of births (size of the 'upward step') occurring in $\mathcal{B}$ ($\mathcal{R}$) if the kth event is a birth event.

A third type of mismatch occurs in coupling the event times in $\mathcal{E}^{(n)}$ and $\mathcal{B}$. Conditionally upon there being no mismatches of the first two types in the first k events and $\eta_{k-1}^{(n)} = \eta_{k-1}$, we have by (4.23) that a mismatch occurs and $\eta_k^{(n)} \neq \eta_k$ only if $U_{k,3} > 1 - \tilde{d}^{(n)}\left(X_{k-1}^{(n)}, Y_{k-1}^{(n)} \right)$.

The following lemma gives conditions under which the processes $\mathcal{B}$ ($\mathcal{R}$) and $\mathcal{E}^{(n)}$ ($\mathcal{S}^{(n)}$) can be constructed on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, so that for $\mathbb{P}$-almost all $\omega \in \Omega$ they coincide over the first u_n events for all sufficiently large n, where $u_n \to \infty$ as $n \to \infty$.

Lemma 4.2. *Suppose that* (3.9) *holds and* (3.10) *holds for some* $\theta_0 > 0$. *Suppose that there exists* $\alpha \geq 1$ *such that* $\mathbb{E}[C^{\alpha+1}] < \infty$, *which in turn implies that* $\mathbb{E}[\tilde{Z}^\alpha] < \infty$.

Let (u_n) *be any non-decreasing sequence of integers such that there exists*

$$0 < \zeta < \min\left\{ \frac{\alpha \theta_0}{2(1+\alpha)}, \frac{\alpha}{2+4\alpha} \right\}, \tag{4.27}$$

so that for all sufficiently large n, $u_n \leq \lfloor K n^\zeta \rfloor$ *for some* $K \in \mathbb{R}^+$.

Then there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, *on which are defined the branching process,* $\mathcal{B}$, *the random walk,* $\mathcal{R}$, *and the sequence of epidemic processes and discrete epidemic processes,* $(\mathcal{E}_n, \mathcal{S}_n)$, *such that for* $\mathbb{P}$-*almost all* $\omega \in \Omega$,

$$\left(Y_1^{(n)}(\omega), Y_2^{(n)}(\omega), \ldots, Y_{u_n}^{(n)}(\omega)\right) = (Y_1(\omega), Y_2(\omega), \ldots, Y_{u_n}(\omega)) \qquad (4.28)$$

and

$$\left(\eta_1^{(n)}(\omega), \eta_2^{(n)}(\omega), \ldots, \eta_{u_n}^{(n)}(\omega)\right) = \left(\eta_1(\omega), \eta_2(\omega), \ldots, \eta_{u_n}(\omega)\right) \qquad (4.29)$$

for all sufficiently large n.

Proof. Without loss of generality, we prove the lemma by taking $u_n = \lfloor Kn^\zeta \rfloor$ for some $K \in (0, \infty)$ and ζ satisfying (4.27). It follows from (4.27) that θ and δ can be chosen so that $\theta, \delta > 0$, $\frac{2(1+\alpha)}{\alpha}\zeta < \theta \leq \theta_0$, and $\theta + 2\zeta + 2\delta < 1$. (Note that (4.27) implies $2\zeta(1+\alpha)/\alpha < \theta_0$. Furthermore,

$$\inf_{\theta > 2(1+\alpha)\zeta/\alpha} \{\theta + 2\zeta\} = \frac{2 + 4\alpha}{\alpha}\zeta < 1,$$

by (4.27).) Set $s_n = n^\theta$, $r_n = Kn^{\zeta+\delta}$, $a_n = \lfloor n^{\theta/(\alpha+1)} \rfloor$, and, for convenience, $\epsilon_n = 1/s_n$. Note that $s_n r_n^2/n \to 0$ as $n \to \infty$, satisfying the conditions of Lemma 4.1.

For $h, n = 1, 2, \ldots$, let $\mathbf{x}_h^{(n)} = \left(x_0^{(n)}, x_1^{(n)}, \ldots, x_h^{(n)}\right)$ and define $\mathbf{y}_h^{(n)}$ similarly. Let

$$\tilde{A}_{n,0} = \left\{ \left(\mathbf{x}_{u_n}^{(n)}, \mathbf{y}_{u_n}^{(n)}\right): \min_{\{1 \leq h \leq u_n\}} x_h^{(n)} = x_{u_n}^{(n)} \geq n - r_n, \max_{\{1 \leq h \leq u_n\}} y_h^{(n)} \leq r_n \right\}$$

and $A_{n,0} = \left\{ \omega \in \Omega : \left(\mathbf{X}_{u_n}^{(n)}(\omega), \mathbf{Y}_{u_n}^{(n)}(\omega)\right) \in \tilde{A}_{n,0} \right\}$. Note that if $\omega \in A_{n,0}$ for all sufficiently large n, then $\left\{ \left(X_k^{(n)}(\omega), Y_k^{(n)}(\omega)\right) \right\}$ satisfies the conditions of Lemma 4.1.

Let H_n denote the event at which the first mismatch occurs between $\mathcal{S}_n$ and $\mathcal{R}$. Then, for $\omega \in \Omega$, (4.28) holds if and only if $H_n(\omega) > u_n$. Note that the first mismatch occurs at event k with $\left(X_{k-1}^{(n)}, Y_{k-1}^{(n)}\right) = \left(x_{k-1}^{(n)}, y_{k-1}^{(n)}\right)$, if

$$U_{h,1} \in D_1^{(n)}\left(x_{k-1}^{(n)}, y_{k-1}^{(n)}\right) \quad \text{or} \quad U_{h,2} \in D_2^{(n)}\left(x_{k-1}^{(n)}, y_{k-1}^{(n)}\right),$$

where $D_1^{(n)}\left(x^{(n)}, y^{(n)}\right)$ and $D_2^{(n)}\left(x^{(n)}, y^{(n)}\right)$ are defined in (4.24) and (4.16), respectively.

Similarly, let $\tilde{H}_n$ denote the event at which the first mismatch occurs between the times of corresponding events in $\mathcal{E}_n$ and $\mathcal{B}$. Then (4.29) holds if and only if $\tilde{H}_n(\omega) > u_n$. Note that if $H_n(\omega) > u_n$ then the first mismatch in the time of events occurs at event k with $\left(X_{k-1}^{(n)}, Y_{k-1}^{(n)}\right) = \left(x_{k-1}^{(n)}, y_{k-1}^{(n)}\right)$, if

$$U_{h,3} \in D_3^{(n)}\left(x_{k-1}^{(n)}, y_{k-1}^{(n)}\right) \equiv \left(1 - \tilde{d}^{(n)}\left(x^{(n)}, y^{(n)}\right), 1\right],$$

where $\tilde{d}^{(n)}\left(x^{(n)}, y^{(n)}\right)$ is defined in (4.19).

By Lemma 4.1, we have for any $\ell > 0$, for all sufficiently large n and $\left(\mathbf{x}_{u_n}^{(n)}, \mathbf{y}_{u_n}^{(n)}\right) \in \tilde{A}_{n,0}$, that $\sum_{w=0}^{\infty} \left| q^{(n)}\left(x^{(n)}, y^{(n)}, w\right) - \beta\varphi_w \right| < \ell\epsilon_n$ and $\left| q^{(n)}\left(x^{(n)}, y^{(n)}\right) - \beta \right| < \ell\epsilon_n$. The first inequality implies that for all $w \in \mathbb{Z}_+$,

$$\left| q^{(n)}\left(x^{(n)}, y^{(n)}, w\right) - \beta\varphi_w \right| < \ell\epsilon_n.$$

Therefore, since $q^{(n)}(x^{(n)}, y^{(n)}) > \beta/2$ for all sufficiently large n, we have by the triangle inequality that

$$
\begin{aligned}
\left| \varpi_w^{(n)}(x^{(n)}, y^{(n)}) - \varphi_w \right| &= \left| \frac{q^{(n)}(x^{(n)}, y^{(n)}, w)}{q^{(n)}(x^{(n)}, y^{(n)})} - \frac{\beta \varphi_w}{\beta} \right| \\
&\leq \frac{1}{q^{(n)}(x^{(n)}, y^{(n)})} \left| q^{(n)}(x^{(n)}, y^{(n)}, w) - \beta \varphi_w \right| \\
&\quad + \beta \varphi_w \left| \frac{1}{q^{(n)}(x^{(n)}, y^{(n)})} - \frac{1}{\beta} \right| \\
&\leq \frac{2}{\beta} \left| q^{(n)}(x^{(n)}, y^{(n)}, w) - \beta \varphi_w \right| + \frac{2\varphi_w}{\beta} \left| q^{(n)}(x^{(n)}, y^{(n)}) - \beta \right| \\
&\leq \frac{4}{\beta} \ell \epsilon_n.
\end{aligned}
$$

Setting $\ell = \frac{\beta}{5}$, we have that for all sufficiently large n, $\left| \varpi_w^{(n)}(x^{(n)}, y^{(n)}) - \varphi_w \right| \leq \epsilon_n$ ($w = 0, 1, \dots$).

Thus we can define sets $\tilde{D}_i^{(n)}$ ($i = 1, 2, 3$) such that for all sufficiently large n, if $(x^{(n)}, y^{(n)}) \in \tilde{A}_n$, then $D_i^{(n)}(x^{(n)}, y^{(n)}) \subseteq \tilde{D}_i^{(n)}$ ($i = 1, 2, 3$), where

$$
\begin{aligned}
\tilde{D}_1^{(n)} &= \left(\frac{\gamma}{\beta + \gamma}, \frac{\gamma}{\beta - \epsilon_n + \gamma} \right], \\
\tilde{D}_2^{(n)} &= \bigcup_{w=0}^{a_n} \left(\sum_{l=0}^{w} \varphi_l - \min\{\varphi_w, \epsilon_n\}, \sum_{l=0}^{w} \varphi_l \right] \cup \left(\sum_{l=0}^{a_n} \varphi_l, 1 \right],
\end{aligned}
$$

and

$$
\tilde{D}_3^{(n)} = (1 - \epsilon_n, 1].
$$

Since ϵ_n is decreasing in n, we have that for all n, $\tilde{D}_i^{(n+1)} \subseteq \tilde{D}_i^{(n)}$ ($i = 1, 2, 3$).

For $i = 1, 2, 3$, let

$$
A_{n,i} = \bigcap_{h=1}^{u_n} \left\{ U_{h,i} \notin \tilde{D}_i^{(n)} \right\}.
$$

We observe that if $u_{n+1} = u_n$, then $A_{n,i} \subseteq A_{n+1,i}$ ($i = 0, 1, 2, 3$). Therefore, following Ball and O'Neill [6, Lemma 2.11], we define $\mathcal{Q} = \{ n \in \mathbb{N} : \lfloor Kn^\zeta \rfloor \neq \lfloor K(n-1)^\zeta \rfloor \}$ and note that, for $i = 0, 1, 2, 3$, to show that

$$
\mathbb{P}\left(A_{n,i}^c \text{ occurs for infinitely many } n \right) = 0,
$$

it is sufficient to show that

$$
\mathbb{P}\left(A_{n,i}^c \text{ occurs for infinitely many } n \in \mathcal{Q} \right) = 0. \tag{4.30}
$$

Given that (4.30) holds for $i = 0, 1, 2, 3$, we have that there exists $\tilde{\Omega} \subseteq \Omega$ such that $\mathbb{P}(\tilde{\Omega}) = 1$ and for every $\omega \in \tilde{\Omega}$, there exists $n(\omega) \in \mathbb{N}$ such that for all $n \geq n(\omega)$, $H_n(\omega) > u_n$ and $\tilde{H}_n(\omega) > u_n$. Thus (4.28) and (4.29) hold, and the lemma follows.

We complete the proof of the lemma by proving (4.30) for $i = 0, 1, 2, 3$. Suppose that, for $i = 0, 1, 2, 3$, there exist $L_i < \infty$ and $\chi_i > 1$ such that, for all sufficiently large n,

$$\mathbb{P}(A_{n,i}^c) \leq L_i n^{-\zeta \chi_i}. \tag{4.31}$$

Following the proof of Ball and O'Neill [6, Lemma 2.10], we have that

$$\sum_{n \in \mathcal{Q}} \mathbb{P}(A_{n,i}^c) \leq \sum_{n \in \mathcal{Q}} L_i n^{-\zeta \chi_i} < \infty,$$

so by the first Borel–Cantelli lemma, (4.30) holds.

Let us prove (4.31). Recall that $\mu_C = \mathbb{E}[C]$, $\mathbb{E}[\tilde{Z}] = \mathbb{E}[C \nu(C)]/\mu_C$, and $\mathbb{E}[C(C-1)\nu(C)] < \infty$, where $\nu(c)$, defined at (3.7), is the mean number of new infectives created in a mixing event of size c with 1 infective and $c-1$ susceptibles. Since $u_n \leq \lfloor Kn^\zeta \rfloor$ and $r_n = Kn^{\zeta + \delta}$, by Chebyshev's inequality, we have that, for all sufficiently large n,

$$\mathbb{P}(A_{n,0}^c) = \mathbb{P}\left(\sum_{j=1}^{u_n} \tilde{Z}_j > r_n - m\right)$$

$$\leq \mathbb{P}\left(\left|\sum_{j=1}^{u_n} \tilde{Z}_j - u_n \mathbb{E}[\tilde{Z}]\right| > \frac{r_n}{2}\right)$$

$$\leq \frac{4}{r_n^2} \mathrm{Var}\left(\sum_{j=1}^{u_n} \tilde{Z}_j\right) \leq \frac{4u_n}{r_n^2} \mathbb{E}[\tilde{Z}_1^2]$$

$$= \frac{4u_n}{r_n^2} \sum_{w=0}^{\infty} w^2 \frac{1}{\mu_C} \sum_{c=w+1}^{\infty} cp_C(c)\pi_c(w; c-1, 1)$$

$$= \frac{4u_n}{r_n^2} \times \frac{1}{\mu_C} \sum_{c=2}^{\infty} cp_C(c) \sum_{w=0}^{c-1} w^2 \pi_c(w; c-1, 1)$$

$$\leq \frac{4u_n}{r_n^2} \times \frac{1}{\mu_C} \sum_{c=2}^{\infty} c(c-1)p_C(c) \sum_{w=0}^{c-1} w\pi_c(w; c-1, 1)$$

$$\leq \frac{4u_n}{r_n^2} \times \frac{\mathbb{E}[C(C-1)\nu(C)]}{\mu_C} \leq \frac{4\mathbb{E}[C(C-1)\nu(C)]}{K\mu_C} n^{-\zeta(1+2\delta/\zeta)}. \tag{4.32}$$

Hence (4.31) holds for $i = 0$.

Since $\theta - 2(1+\alpha)\zeta/\alpha > 0$, we have that for all sufficiently large n,

$$\mathbb{P}(A_{n,1}^c) = \mathbb{P}\left(\bigcup_{h=1}^{u_n} \left\{U_{h,1} \in \tilde{D}_1^{(n)}\right\}\right)$$

$$\leq \left(\frac{2\gamma}{(\beta+\gamma)^2}\right) u_n \epsilon_n$$

$$\leq \left(\frac{2\gamma}{(\beta+\gamma)^2}\right) Kn^\zeta n^{-\theta}$$

$$\leq \left(\frac{2\gamma}{(\beta+\gamma)^2}\right) Kn^{-\zeta(1+2/\alpha)}. \tag{4.33}$$

Hence (4.31) holds for $i = 1$.

Similarly, since $\mathbb{P}(A_{n,3}^c) \le u_n \epsilon_n$, we have that (4.31) holds for $i = 3$.

Finally, let $\delta_1 = \frac{\alpha}{\zeta(1+\alpha)}\theta - 2 > 0$. For all sufficiently large n, we have that a_n^α, $s_n/a_n \ge \frac{1}{2}n^{\theta\alpha/(1+\alpha)}$. Thus, recalling that $\epsilon_n = 1/s_n$, we have that for all sufficiently large n,

$$\mathbb{P}(A_{n,2}^c) = \mathbb{P}\left(\bigcup_{h=1}^{u_n} \left\{U_{h,2} \in \tilde{D}_2^{(n)}\right\}\right)$$

$$\le u_n \left\{a_n \epsilon_n + \mathbb{P}(\tilde{Z} > a_n)\right\}$$

$$\le u_n \left\{\frac{a_n}{s_n} + \frac{\mathbb{E}[\tilde{Z}^\alpha]}{a_n^\alpha}\right\}$$

$$\le 2(1 + \mathbb{E}[\tilde{Z}^\alpha])Kn^\zeta n^{-\theta\alpha/(1+\alpha)}$$

$$= 2(1 + \mathbb{E}[\tilde{Z}^\alpha])n^{-\zeta(1+\delta_1)}. \tag{4.34}$$

Hence (4.31) holds for $i = 2$. Thus (4.30) holds for $i = 0, 1, 2, 3$ and the lemma is proved. $\quad\square$

Lemma 4.2 ensures that the the processes $\mathcal{E}^{(n)}$ ($\mathcal{S}^{(n)}$) and $\mathcal{B}$ ($\mathcal{R}$) coincide for an increasing number of events as n increases. For Theorem 3.1(a) we do not require as strong a result as Lemma 4.2, and the following corollary, which can be proved in a similar fashion to Lemma 4.2, suffices.

Corollary 4.1. *For any $K \in \mathbb{N}$, we have, for $(\Omega, \mathcal{F}, \mathbb{P})$ defined in Lemma 4.2, that for $\mathbb{P}$-almost all $\omega \in \Omega$,*

$$\left(Y_1^{(n)}(\omega), Y_2^{(n)}(\omega), \ldots, Y_K^{(n)}(\omega)\right) = (Y_1(\omega), Y_2(\omega), \ldots, Y_K(\omega))$$

and

$$\left(\eta_1^{(n)}(\omega), \eta_2^{(n)}(\omega), \ldots, \eta_K^{(n)}(\omega)\right) = (\eta_1(\omega), \eta_2(\omega), \ldots, \eta_K(\omega)),$$

for all sufficiently large n.

The coupling in Lemma 4.2 includes birth events where no births occur, that is, $Z_1 = 0$. Given that $\lambda < \infty$ ($\beta = \lambda \mu_C < \infty$) and $\gamma > 0$, it follows that $\mathbb{P}(Z_1 \ne 0) > 0$. Since $Z_1, Z_2, \ldots$ are i.i.d., the strong law of large numbers yields

$$\frac{1}{p}\sum_{i=1}^{p} 1_{\{Z_i \ne 0\}} \xrightarrow{\text{a.s.}} \mathbb{P}(Z_1 \ne 0) \qquad \text{as } p \to \infty, \tag{4.35}$$

where $\xrightarrow{\text{a.s.}}$ denotes convergence almost surely. For $k = 1, 2, \ldots$, let

$$M_k = \min\left\{p: \sum_{i=1}^{p} 1_{\{Z_i \ne 0\}} = k\right\}. \tag{4.36}$$

Thus M_k is the kth event in $\mathcal{B}$ for which $Z_i \ne 0$. (If $\mathcal{B}$ goes extinct then M_k has this interpretation for only finitely many k.) Theorem 3.1 now follows straightforwardly from Lemma 4.2.

Proof of Theorem 3.1. (a) Recall that T is the total size of the branching process $\mathcal{B}$ and $A_{ext} = \{\omega \in \Omega : T(\omega) < \infty\}$.

Fix $\omega \in A_{ext}$ and suppose that $T(\omega) = k < \infty$. Then there exists $h = h(\omega) \le 2k - m$ such that $Y_{M_h}(\omega) = 0$. That is, there are at most $k - m$ birth events (with a strictly positive number

of births) and k death events in the branching process. By Corollary 4.1 there exists $n_2(\omega) \in \mathbb{N}$ such that for all $n \geq n_2(\omega)$ and $l = 1, 2, \ldots, M_h(\omega)$, $Y_l^{(n)}(\omega) = Y_l(\omega)$ and $\eta_l^{(n)}(\omega) = \eta_l(\omega)$, and hence, for all $t \geq 0$, $I_n(t, \omega) = I(t, \omega)$.

(b) Let ρ satisfy (3.11) and $t_n = \rho \log n$. Remembering from (3.8) that $r = \gamma(R_0 - 1)$ is the Malthusian parameter (growth rate) of the branching process, we take ζ such that

$$\rho r < \zeta < \min \left\{ \frac{\alpha \theta_0}{2(1+\alpha)}, \frac{\alpha}{2+4\alpha} \right\},$$

so that ζ satisfies (4.27) in the statement of Lemma 4.2.

For $t \geq 0$, let $N(t)$ denote the total number of (birth and death) events in the branching process $\mathcal{B}$ up to and including time t. Then, if $N(t_n, \omega) \leq u_n = \lfloor n^\zeta \rfloor$ and $\left(Y_h^{(n)}(\omega), \eta_h^{(n)}(\omega) \right) = (Y_h(\omega), \eta_h(\omega))$ $(h = 1, 2, \ldots, u_n)$, we have from Lemma 4.2 that

$$\sup_{0 \leq t \leq t_n} \left| I^{(n)}(t) - I(t) \right| = 0.$$

Give the initial ancestors the labels $-(m-1), -(m-2), \ldots, 0$, and label the individuals born in the branching process sequentially $1, 2, \ldots$. For $i = 1, 2, \ldots$, let τ_i denote the time of the birth of the ith individual, with the conventions that $\tau_i = \infty$ if fewer than i births occur, and $\tau_i = 0$ for $i = -(m-1), -(m-2), \ldots, 0$. For $i = -(m-1), -(m-2), \ldots$, let $\tilde{G}_i(s)$ denote the number of birth and death events involving individual i in the first s time units after their birth, if $s \geq 0$, and let $\tilde{G}_i(s) = 0$ if $s < 0$. Note that $\tilde{G}_i(s)$ is non-decreasing in s and $\tilde{G}_i(\infty) \stackrel{D}{=} G + 1$, where G is the number of birth events involving an individual and is a geometric random variable given by (3.4). Therefore, for all $t \geq 0$,

$$N(t) = \sum_{i=-(m-1)}^{\infty} \tilde{G}_i(t - \tau_i). \tag{4.37}$$

Note that $N(t)$ satisfies the form of Nerman [14, (1.11)]. It is straightforward to show that the conditions of Theorem 5.4 in Nerman [14] hold, since $\mathbb{E}[\tilde{G}(\infty)] = (\gamma + \lambda \mu_C)/\gamma$. Therefore, by that theorem, there exists a positive, almost surely finite random variable W such that

$$\lim_{t \to \infty} \exp(-rt) N(t, \omega) = W(\omega) > 0$$

for $\mathbb{P}$-almost all $\omega \in A_{\text{ext}}^c$. It is then straightforward to show, following the proof of Ball and O'Neill [6, Lemma 2.9], that there exists some $\mathbb{P}$-measurable set $B_1 \subseteq A_{\text{ext}}^c$ such that $\mathbb{P}(B_1) = \mathbb{P}(A_{\text{ext}}^c)$ and for all $\omega \in B_1$,

$$\lim_{t \to \infty} n^{-cr} N(t_n, \omega) = W(\omega).$$

Hence, for all sufficiently large n, $N(t_n, \omega) \leq 2W(\omega) n^{\rho r} \leq u_n$. Finally, by Lemma 4.2, for $\mathbb{P}$-almost all $\omega \in B_1$, (4.28) and (4.29) hold, so (3.12) follows. $\qquad \square$

Finally, we consider the total size of the epidemic processes and branching processes, with Corollary 3.1 following straightforwardly from Corollary 4.1.

Proof of Corollary 3.1. Let $\hat{\Omega} \subseteq \Omega$ be the set on which the convergence underlying (4.35) holds, so $\mathbb{P}(\hat{\Omega}) = 1$, and fix $\omega \in \tilde{\Omega} \cap \hat{\Omega}$. Suppose that $T(\omega) = k < \infty$. Then there exists

$h = h(\omega) \leq 2k - m$ such that $Y_{M_h}(\omega) = 0$. By Corollary 4.1, there exists $\tilde{n}(\omega) \in \mathbb{N}$ such that for all $n \geq \tilde{n}(\omega)$, $Y_i^{(n)}(\omega) = Y_i(\omega)$ $(i = 1, 2, \ldots, M_h(\omega))$. Thus, $T^{(n)}(\omega) = T(\omega)$ for all $n \geq \tilde{n}(\omega)$.

Suppose instead that $T(\omega) = \infty$. Choose any $k_1 \in \mathbb{N}$ and let $n_E(k_1, \omega)$ be the number of events in $\mathcal{B}$ when the total size of $\mathcal{B}$ first reaches at least k_1. Then arguing as above, with k replaced by k_1 and h replaced by $n_E(k_1, \omega)$, shows that $T^{(n)}(\omega) \geq k_1$ for all sufficiently large n. This holds for all $k_1 \in \mathbb{N}$, so $T^{(n)}(\omega) \to \infty$ as $n \to \infty$. □

5. Proof of Theorem 3.2

5.1. Overview

We present an overview of the steps to prove Theorem 3.2. In Section 5.2, we prove Theorem 3.2(a), which states that as $n \to \infty$, the probability of a major epidemic outbreak in $\mathcal{E}^{(n)}$ (the epidemic infects at least $\log n$ individuals) tends to the probability that the branching process, $\mathcal{B}$, does not go extinct. In Section 5.3, we introduce a sequence of lower-bound random walks, $(\mathcal{L}^{(n)})$, which is a key component in showing that a major epidemic in the discrete epidemic jump process, $\mathcal{S}^{(n)}$, and hence in the epidemic process $\mathcal{E}^{(n)}$ infects at least $\delta^* n$ individuals with probability tending to 1 as $n \to \infty$. We provide an outline of the coupling of $\mathcal{S}^{(n)}$ and $\mathcal{L}^{(n)}$, via an intermediary process $\mathcal{G}^{(n)}$, and in (5.5) we identify the relationship between the three processes with $\mathcal{L}^{(n)}$ as a lower bound in terms of the number of infectives in the epidemic, to establish Theorem 3.2(b). The details of $\mathcal{L}^{(n)}$ are provided in Section 5.4, along with Lemmas 5.1 and 5.2, which provide the main steps in establishing (5.5). Finally, Section 5.4 concludes with the proof of Theorem 3.2(b), from which Corollary 3.2 follows immediately.

5.2. Probability of a major epidemic

Under the conditions of Theorem 3.2, the conditions of Lemma 4.2 are satisfied with $\alpha = 1$ since $\mathbb{E}[C^2] < \infty$. The proof of Theorem 3.2(a) then follows almost immediately from the proof of Theorem 3.1 by considering the embedded random walk and discrete-time epidemic jump process. From Lemma 4.2, (4.28), we have that $\mathbf{Y}_{u_n} = (Y_1, Y_2, \ldots, Y_{u_n})$ and $\mathbf{Y}_{u_n}^{(n)} = \left(Y_1^{(n)}, Y_2^{(n)}, \ldots, Y_{u_n}^{(n)}\right)$ can be constructed so that

$$\mathbb{P}\left(\mathbf{Y}_{u_n}^{(n)} = \mathbf{Y}_{u_n}\right) \to 1 \quad \text{as } n \to \infty, \tag{5.1}$$

for $u_n = \lfloor n^\zeta \rfloor$ and $\zeta > 0$ satisfying (4.27) with $\alpha = 1$. Hence we can couple the process $\mathcal{S}^{(n)}$ to $\mathcal{R}$ over the first u_n steps, and Theorem 3.2(a) follows, as we now show.

Proof of Theorem 3.2(a). Since T is the total size of a branching process with m initial ancestors, it follows that

$$\mathbb{P}(T \geq \log n) \to 1 - z^m \quad \text{as } n \to \infty. \tag{5.2}$$

Following the proof of Corollary 3.1, $T < \log n$ if there exists $h_n \leq 2 \log n - m$ such that $Y_{M_{h_n}} = 0$. Let $\tilde{\Omega}$ and $\hat{\Omega}$ be as in the proofs of Lemma 4.2 and Corollary 3.1, respectively, and note that $\mathbb{P}(\tilde{\Omega} \cap \hat{\Omega}) = 1$. Fix $\omega \in \tilde{\Omega} \cap \hat{\Omega}$. Then

$$M_{\lfloor 2 \log n \rfloor}(\omega) \leq \frac{3}{\mathbb{P}(Z_1 \neq 0)} \log n \leq u_n$$

for all sufficiently large n, where M_k is given in (4.36) and $\tilde{\Omega}$ and $\hat{\Omega}$ are defined in Lemma 4.2 and Corollary 3.1, respectively, with $\mathbb{P}(\tilde{\Omega} \cap \hat{\Omega}) = 1$. It follows using Lemma 4.2 that $1_{\{T(\omega) < \log n\}} = 1_{\{T^{(n)}(\omega) < \log n\}}$ for all sufficiently large n. Thus, $1_{\{T < \log n\}} - 1_{\{T^{(n)} < \log n\}}$ converges almost surely (and hence in probability) to 0 as $n \to \infty$. Therefore,

$$\left| \mathbb{P}\big(T^{(n)} \geq \log n\big) - \mathbb{P}(T \geq \log n) \right| = \left| \mathbb{P}\big(T^{(n)} < \log n\big) - \mathbb{P}(T < \log n) \right| \to 0 \quad \text{as } n \to \infty,$$

and (3.13) follows immediately using (5.2). □

5.3. Coupling of the lower-bound random walk to the epidemic

We turn to the proof of Theorem 3.2(b) and note that we now need to consider the epidemic process and any approximation over $\lfloor \delta_1 n \rfloor$ events for some $\delta_1 > 0$. The couplings utilised thus far do not extend to $\lfloor \delta_1 n \rfloor$ events in the limit as $n \to \infty$. However, we can still utilise the couplings over the first $u_n = \lfloor n^\zeta \rfloor$ events. Hence, given that the embedded discrete epidemic jump process $\mathcal{S}^{(n)}$ reaches u_n events without hitting 0 (the epidemic process $\mathcal{E}^{(n)}$ does not go extinct), we can show, following the proof of Theorem 3.2(a), that, with probability tending to 1 as $n \to \infty$, $T^{(n)} \geq v_n$, where $v_n = \mathbb{P}(Z_1 \neq 0)u_n/3$. It immediately follows using the coupling in Lemma 4.2 (see (5.1)) that $\mathbb{P}\big(T \geq v_n | T^{(n)} \geq v_n\big) \to 1$ as $n \to \infty$. We have that if $T \geq v_n$, then $\min_{k \leq v_n}\{Y_k\} > 0$, and by the weak law of large numbers,

$$\left| \frac{1}{v_n} \sum_{k=1}^{v_n} Z_k - \mathbb{E}[Z] \right| \xrightarrow{\text{p}} 0 \quad \text{as } n \to \infty.$$

Hence, under the assumption $R_0 > 1$, which is required for Theorem 3.2(b), we have that $\mathbb{E}[Z] > 0$ and, since $Y_{v_n} = m + \sum_{k=1}^{v_n} Z_k$, that

$$\mathbb{P}\left(Y_{v_n} > \frac{v_n \mathbb{E}[Z]}{2}, \min_{k \leq v_n}\{Y_k\} > 0 \,\middle|\, T \geq v_n \right) \to 1 \quad \text{as } n \to \infty. \tag{5.3}$$

Now (5.3) implies that if the branching process $\mathcal{B}$ has at least v_n individuals ever alive, then the number of individuals alive in $\mathcal{B}$ (the position of the random walk $\mathcal{R}$) after v_n events exceeds $v_n \mathbb{E}[Z]/2$ with probability tending to 1 as $n \to \infty$. Combined with (5.1) the same holds true for $Y_{v_n}^{(n)}$ in $\mathcal{E}^{(n)}$ and $\mathcal{S}^{(n)}$.

The next step is to show that, given that the epidemic $\mathcal{E}^{(n)}$ ($\mathcal{S}^{(n)}$) has not gone extinct in v_n events, there exists $\delta^* > 0$ such that, with probability tending to 1 as $n \to \infty$, at least $\lfloor \delta^* n \rfloor$ events occur in $\mathcal{S}^{(n)}$. In order to do this, we introduce a lower-bound random walk $\mathcal{L}^{(n)}$ indexed by the population size n. Lower-bound branching processes (random walks) for epidemic processes go back to [17], and the main idea is along similar lines to [17], in that we set up the lower-bound random walk so that the number of infectives in the discrete epidemic jump process $\mathcal{S}^{(n)}$ is at least the number of individuals alive in the branching process with embedded random walk $\mathcal{L}^{(n)}$ for the initial stages of the epidemic process.

The key features in setting up $\mathcal{L}^{(n)}$ are as follows. Let $L_k^{(n)}$ denote the position of the random walk $\mathcal{L}^{(n)}$ after k steps. The random walk $\mathcal{L}^{(n)}$ is set identical to the random walk $\mathcal{R}$ for the first v_n steps; that is, the distribution of steps is according to Z given in (4.1). Hence $L_0^{(n)} = m$, and for $k = 1, 2, \ldots, v_n, L_k^{(n)} = Y_k$. For $k = v_n + 1, v_n + 2, \ldots$, the steps in $\mathcal{L}^{(n)}$ are i.i.d. according to $\hat{Z}^{(n)}$ defined below in (5.9), with $\mathbb{E}[\hat{Z}^{(n)}] > 0$ so that the lower-bound random walk has positive drift. Therefore we can show that, as $n \to \infty$, if $\mathcal{L}^{(n)}$ has not hit 0 in the first v_n steps

when it is coupled to $\mathcal{R}$, and hence reached $L_{v_n}^{(n)} = Y_{v_n} \geq v_n \mathbb{E}[Z]/2$ (cf. (5.3)), with probability tending to 1 it will not hit 0 in the first $\lfloor \delta_1 n \rfloor$ steps for $\delta_1 > 0$.

It is difficult to directly couple $\mathcal{S}^{(n)}$ and $\mathcal{L}^{(n)}$, owing to differences in the distribution of steps caused by the changing rate of events in $\mathcal{E}^{(n)}$ and hence the probability of events occurring in $\mathcal{S}^{(n)}$. Therefore we introduce an intermediary process $\mathcal{G}^{(n)}$. The intermediary process $\mathcal{G}^{(n)}$ is a bivariate (epidemic) process, indexed by the population size n, whose steps are state-dependent, with the dependence corresponding to the number of susceptibles and infectives in the population. For $k = 1, 2, \ldots$, let $\left(A_k^{(n)}, G_k^{(n)} \right)$ denote the state of $\mathcal{G}^{(n)}$ after k steps (events), with $A_k^{(n)}$ and $G_k^{(n)}$ denoting the numbers of susceptibles and infectives, respectively. For the first v_n steps, $\mathcal{G}^{(n)}$ is set identical to $\mathcal{S}^{(n)}$, so that for $k = 1, 2, \ldots, v_n$, $\left(A_k^{(n)}, G_k^{(n)} \right) = \left(X_k^{(n)}, Y_k^{(n)} \right)$. After v_n steps have occurred in both $\mathcal{G}^{(n)}$ and $\mathcal{S}^{(n)}$, we allow the two processes to differ as follows. The process $\mathcal{G}^{(n)}$ is associated with an epidemic-type process, $\mathcal{E}_G^{(n)}$, which has a higher rate of events than the epidemic process $\mathcal{E}^{(n)}$, but in such a way that the additional events in $\mathcal{E}_G^{(n)}$, which do not occur in $\mathcal{E}^{(n)}$, are infection events where no infections occur. In this way we can construct $\mathcal{E}_G^{(n)}$ from $\mathcal{E}^{(n)}$ so that all events in $\mathcal{E}^{(n)}$ occur in $\mathcal{E}_G^{(n)}$, but there are additional *ghost* events which occur in $\mathcal{E}_G^{(n)}$ where there is no change in state (no infection or removal occurs); the only change is to increment the counter of the number of events. Similarly, we can reverse this process and generate $\mathcal{E}^{(n)}$ from $\mathcal{E}_G^{(n)}$ by eliminating, with an appropriate probability, some of the events where no change in state occurs. Therefore, for $k = 1, 2, \ldots$, there exists $\kappa_n(k) \leq k$ such that

$$\left(X_{\kappa_n(k)}^{(n)}, Y_{\kappa_n(k)}^{(n)} \right) = \left(A_k^{(n)}, G_k^{(n)} \right). \tag{5.4}$$

Note that for $k = 1, 2, \ldots, v_n$, $\kappa_n(k) = k$.

We couple $\mathcal{G}^{(n)}$ and $\mathcal{L}^{(n)}$ so that, with probability tending to 1 as $n \to \infty$, we have $G_k^{(n)} \geq L_k^{(n)}$ for $k = 1, 2, \ldots, \lfloor \delta_1 n \rfloor$, where $\delta_1 > 0$ is given in (5.7). That is, for the first $\lfloor \delta_1 n \rfloor$ events, the number of infectives in the process $\mathcal{G}^{(n)}$ is at least the number of individuals alive in the random walk $\mathcal{L}^{(n)}$. It then follows that, for $k = 1, 2, \ldots, \lfloor \delta_1 n \rfloor$, we have

$$Y_{\kappa_n(k)}^{(n)} = G_k^n \geq L_k^{(n)} \tag{5.5}$$

with probability tending to 1 as $n \to \infty$. The proof of Theorem 3.2(b) follows almost immediately after we have established (5.5).

5.4. Lower bound for the size of a major epidemic outbreak

In this section we formally define the lower-bound random walk $\mathcal{L}^{(n)}$. Then, in Lemma 5.1, we show that $\mathcal{L}^{(n)}$ has positive drift, so that after $\lfloor \delta_1 n \rfloor$ steps we have, for some $\delta > 0$ (defined in (5.10)), that $L_{\lfloor \delta_1 n \rfloor}^{(n)} \geq \delta n$ with probability tending to 1 as $n \to \infty$. This is followed by the construction of $\mathcal{G}^{(n)}$. We show that whilst fewer than $\lfloor \delta_1 n \rfloor$ events have occurred, the number of susceptibles (with high probability) remains above $(1 - \epsilon_1)n$, where $\epsilon_1 > 0$ is given in (5.7), which enables us to show, in Lemma 5.2, that the inequality in (5.5) holds with probability tending to 1 as $n \to \infty$. We then establish the equality in (5.5) through the coupling of $\mathcal{S}^{(n)}$ and $\mathcal{G}^{(n)}$, with the proof of Theorem 3.2(b) following.

By (4.18), we have, for all $\left(x^{(n)}, y^{(n)} \right)$, that $q^{(n)}\left(x^{(n)}, y^{(n)} \right) \leq \lambda \mu_C^{(n)} = \beta^{(n)}$, say. Since $R_0 > 1$, and by (3.9), $\mathbb{E}[C(C-1)v(C)] < \infty$, we can define

$$\epsilon = \frac{(R_0 - 1)\gamma}{2\lambda \mathbb{E}[C(C-1)v(C)]} > 0. \tag{5.6}$$

We can then fix δ_1 and $\epsilon_1 < \epsilon$ such that

$$0 < \delta_1 < \epsilon_1 \frac{\mu_C}{\mathbb{E}[C^2]}. \tag{5.7}$$

Throughout the remainder of the section we assume that δ_1 and ϵ_1 satisfy (5.7). For $n = 1, 2, \ldots$ and $w = 1, 2, \ldots, n - 1$, let

$$\psi_w^{(n)} = \max \left\{ \frac{\lambda}{\beta^{(n)}} \sum_{c=w+1}^{n} c p_C^{(n)}(c) \pi_c(w; c - 1, 1)[1 - 2\epsilon_1(c - 1)], 0 \right\} \tag{5.8}$$

and $\psi_0^{(n)} = 1 - \sum_{v=1}^{n-1} \psi_v^{(n)}$. Note that there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $2\epsilon_1 \lfloor n/2 \rfloor > 1$, and hence $\psi_w^{(n)} = 0$ for all $w > \lfloor n/2 \rfloor$. We define

$$\mathbb{P}(\hat{Z}^{(n)} = w) = \begin{cases} \dfrac{\beta^{(n)} \psi_w^{(n)}}{\gamma + \beta^{(n)}}, & w = 1, 2, \ldots, n - 1, \\[2mm] \dfrac{\beta^{(n)}}{\gamma + \beta^{(n)}} \left(1 - \sum_{i=1}^{n-1} \psi_i^{(n)} \right), & w = 0, \\[2mm] \dfrac{\gamma}{\gamma + \beta^{(n)}}, & w = -1, \\[2mm] 0 & \text{otherwise.} \end{cases} \tag{5.9}$$

Let $\hat{Z}_{v_n+1}^{(n)}, \hat{Z}_{v_n+2}^{(n)}, \ldots$, be i.i.d. according to $\hat{Z}^{(n)}$ given in (5.9). Then, given $L_{v_n}^{(n)} = Y_{v_n}$, for $v_n + 1 \leq k \leq \lfloor \delta_1 n \rfloor$ we set

$$L_k^{(n)} = L_{k-1}^{(n)} + \hat{Z}_k^{(n)}.$$

Lemma 5.1. *Let b_n be any sequence of positive integers such that $b_n \to \infty$ as $n \to \infty$. Let δ_1 satisfy (5.7), and let δ satisfy*

$$0 < \delta < \delta_1 \frac{(R_0 - 1)\gamma - 2\epsilon_1 \lambda \mathbb{E}[C(C - 1)\nu(C)]}{\gamma + \lambda \mu_C}, \tag{5.10}$$

where $\epsilon_1 < \epsilon$ ensures that the right-hand side of (5.10) is positive. Then

$$\mathbb{P} \left(L_{\lfloor \delta_1 n \rfloor}^{(n)} \geq \lfloor \delta n \rfloor, \min_{k > v_n} \left\{ L_k^{(n)} > 0 \right\} \, \Big| \, L_{v_n}^{(n)} \geq b_n \right) \to 1 \quad \text{as } n \to \infty, \tag{5.11}$$

where $v_n = \mathbb{P}(Z_1 \neq 0) u_n / 3$.

Proof. Using (5.9) and (5.8),

$$\mathbb{E}[\hat{Z}^{(n)}] = -\frac{\gamma}{\gamma + \beta^{(n)}} + \sum_{w=1}^{n-1} w \frac{\beta^{(n)} \psi_w^n}{\gamma + \beta^{(n)}}$$

$$\geq \frac{1}{\gamma + \beta^{(n)}} \left\{ -\gamma + \sum_{w=1}^{n-1} w \left[\lambda \sum_{c=w+1}^{n} c p_C^{(n)}(c) \pi_c(w; c - 1, 1) [1 - 2\epsilon_1(c - 1)] \right] \right\}$$

$$= \frac{1}{\gamma + \beta^{(n)}} \left\{ -\gamma + \lambda \sum_{c=2}^{n} c p_C^{(n)}(c) \sum_{w=1}^{c-1} w \pi_c(w; c-1, 1) \right.$$

$$\left. -2\epsilon_1 \lambda \sum_{c=2}^{n} c(c-1) p_C^{(n)} \sum_{w=1}^{c-1} w \pi_c(w; c-1, 1) \right\}$$

$$= \frac{1}{\gamma + \beta^{(n)}} \left\{ -\gamma + \lambda \mathbb{E}[C^{(n)} \nu(C^{(n)})] - 2\epsilon_1 \lambda \mathbb{E}[C^{(n)}(C^{(n)} - 1) \nu(C^{(n)})] \right\}.$$

Recall the expression (3.6) for R_0 and the fact that by (3.9), $\mathbb{E}[C(C-1)\nu(C)] < \infty$. Since $\lambda \mathbb{E}[C^{(n)} \nu(C^{(n)})] \to R_0 \gamma$ and $\beta^{(n)} \to \beta = \lambda \mu_C$ as $n \to \infty$, we have that

$$\liminf_{n \to \infty} \mathbb{E}[\hat{Z}^{(n)}] \geq \frac{(R_0 - 1)\gamma - 2\epsilon_1 \lambda \mathbb{E}[C(C-1)\nu(C)]}{\gamma + \lambda \mu_C} > 0,$$

where the final inequality follows from (5.10). It also follows from (5.8) that $\beta^{(n)} \psi_w^{(n)} \leq \lambda \sum_{c=w+1}^{n} c p_C^{(n)}(c) \pi_c(w; c-1, 1)$, for all $w = 1, 2, \ldots, n-1$, so

$$\mathbb{E}[(\hat{Z}^{(n)})^2] \leq \frac{\gamma}{\gamma + \beta^{(n)}} + \frac{1}{\gamma + \beta^{(n)}} \sum_{w=1}^{n-1} w^2 \lambda \sum_{c=w+1}^{n} c p_C^{(n)}(c) \pi_c(w; c-1, 1)$$

$$= \frac{1}{\gamma + \lambda \mathbb{E}[C^{(n)}]} \left\{ \gamma + \lambda \sum_{c=2}^{n} c p_C^{(n)}(c) \sum_{w=1}^{c-1} w^2 \pi_c(w; c-1, 1) \right\}$$

$$\leq \frac{1}{\gamma + \lambda \mathbb{E}[C^{(n)}]} \left\{ \gamma + \lambda \sum_{c=2}^{n} c(c-1) p_C^{(n)}(c) \sum_{w=1}^{c-1} w \pi_c(w; c-1, 1) \right\}$$

$$= \frac{1}{\gamma + \lambda \mathbb{E}[C^{(n)}]} \left\{ \gamma + \lambda \mathbb{E}[C^{(n)}(C^{(n)} - 1) \nu(C^{(n)})] \right\}$$

$$\to \frac{\gamma + \lambda \mathbb{E}[C(C-1)\nu(C)]}{\gamma + \lambda \mu_C} \qquad \text{as } n \to \infty, \tag{5.12}$$

where $\mathbb{E}[C(C-1)\nu(C)] < \infty$ ensures that the right-hand side of (5.12) is finite.

Let z_n be the probability that the random walk $\mathcal{L}^{(n)}$ ever hits 0 given $L_{\nu_n}^{(n)} = 1$. Since $\liminf_{n \to \infty} \mathbb{E}[\hat{Z}^{(n)}] > 0$ and $\sup_n \mathbb{E}[(\hat{Z}^{(n)})^2] < \infty$, it follows from Ball and Neal [4, Lemma A.3] that $\limsup_{n \to \infty} z_n < 1$. Therefore, for any $b_n \to \infty$ as $n \to \infty$, we have that $z_n^{b_n} \to 0$ as $n \to \infty$. Moreover, for $\sup_n \mathbb{E}[(\hat{Z}^{(n)})^2] < \infty$, it follows by the weak law of large numbers for triangular arrays (e.g. Durrett [9, Theorem 2.2.4]) that

$$\left| \frac{1}{\lfloor \delta_1 n \rfloor - \nu_n} \sum_{l=\nu_n+1}^{\lfloor \delta_1 n \rfloor} \hat{Z}_l^n - \mathbb{E}[\hat{Z}^n] \right| \xrightarrow{\text{p}} 0 \quad \text{as } n \to \infty.$$

Hence, for any $b_n \to \infty$ and δ satisfying (5.10), we have that (5.11) holds. □

Turning to the intermediary (ghost) process $\mathcal{G}^{(n)}$, we define independent random variables $W_k^{(n)}(x^{(n)}, y^{(n)})$ ($k = \nu_n + 1, \nu_n + 2, \ldots$) to define the transitions given the current state $(A^{(n)}, G^{(n)}) = (x^{(n)}, y^{(n)})$ after event ν_n. For $k = \nu_n + 1, \nu_n + 2, \ldots$, let $W_k^{(n)}(x^{(n)}, y^{(n)})$ satisfy

$$
\mathbb{P}\big(W_k^{(n)}\big(x^{(n)}, y^{(n)}\big) = w\big) = \begin{cases} \dfrac{q^{(n)}(x^{(n)}, y^{(n)}, w)}{\gamma + \beta^{(n)}}, & w = 1, 2, \ldots, n-1, \\[2ex] \dfrac{\beta^{(n)} - \sum_{v=1}^{n-1} q^{(n)}(x^{(n)}, y^{(n)}, v)}{\gamma + \beta^{(n)}}, & w = 0, \\[2ex] \dfrac{\gamma}{\gamma + \beta^{(n)}}, & w = -1, \\[1ex] 0 & \text{otherwise.} \end{cases} \tag{5.13}
$$

Then for $\big(A_{k-1}^{(n)}, G_{k-1}^{(n)}\big) = \big(x^{(n)}, y^{(n)}\big)$, we set

$$
\big((A_k^{(n)}, G_k^{(n)}\big) = \bigg(x^{(n)} - W_k^{(n)}\big(x^{(n)}, y^{(n)}\big)\mathbf{1}_{\big\{W_k^{(n)}(x^{(n)}, y^{(n)}) \geq 0\big\}}, \; y^{(n)} + W_k^{(n)}\big(x^{(n)}, y^{(n)}\big)\bigg).
$$

The continuous-time epidemic-type process $\mathcal{E}_G^{(n)}$ is constructed from $\mathcal{G}^{(n)}$ as follows. If $G_k^{(n)} = y^{(n)}$, then the time from the kth to the $(k+1)$th event is drawn from $\mathrm{Exp}\big((\gamma + \beta^{(n)})y^{(n)}\big)$, regardless of $A^{(n)}$, the number of susceptibles. Therefore, if there are $y^{(n)}$ infectives in the population, mixing events occur at rate $\beta^{(n)}y^{(n)}$, with the number of individuals infected in such a mixing event depending on the number of susceptibles, $A^{(n)}$.

We consider the coupling of $\mathcal{G}^{(n)}$ and $\mathcal{L}^{(n)}$ in Lemma 5.2 before finalising the coupling between $\mathcal{G}^{(n)}$ and $\mathcal{S}^{(n)}$.

Lemma 5.2. *There exists a coupling of $\mathcal{G}^{(n)}$ and $\mathcal{L}^{(n)}$ such that, for any δ_1 satisfying* (5.7),

$$
\mathbb{P}\left(\bigcap_{k=1}^{\lfloor \delta_1 n \rfloor} \big\{G_k^{(n)} \geq L_k^{(n)}\big\}\right) \to 1 \quad \text{as } n \to \infty. \tag{5.14}
$$

Proof. By the construction of $\mathcal{G}^{(n)}$ and $\mathcal{L}^{(n)}$ and Lemma 4.2, (4.28), we have that with probability tending to 1, $G_k^{(n)} = Y_k = L_k^{(n)}$ ($k = 1, 2, \ldots, \nu_n$).

The first step is to show that $\mathbb{P}\big(A_{\lfloor \delta_1 n \rfloor}^{(n)} \geq (1 - \epsilon_1)n\big) \to 1$ as $n \to \infty$, which, since $A_k^{(n)}$ is non-increasing in k, implies that, for all $k = 1, 2, \ldots, \lfloor \delta_1 n \rfloor$, $A_k^{(n)} \geq (1 - \epsilon_1)n$ with probability tending to 1 as $n \to \infty$.

It is straightforward to show, for all $\big(x^{(n)}, y^{(n)}\big)$, that $W^{(n)}\big(x^{(n)}, y^{(n)}\big)$ is stochastically smaller than $(\overset{st}{\leq})$ a random variable $\tilde{W}^{(n)}$, where $\mathbb{P}(\tilde{W}^{(n)} = -1) = \gamma/\big(\gamma + \beta^{(n)}\big)$ and for $k = 1, 2, \ldots$, $\mathbb{P}(\tilde{W}^{(n)} = k) = \beta^{(n)}\mathbb{P}(\tilde{C}^{(n)} = k + 1)/\big(\gamma + \beta^{(n)}\big)$, with

$$
\mathbb{P}(\tilde{C}^{(n)} = c) = \begin{cases} \dfrac{c p_C^{(n)}(c)}{\mu_C^{(n)}}, & c = 2, 3, \ldots, \\[1ex] 0 & \text{otherwise.} \end{cases}
$$

Note that $\tilde{C}^{(n)}$ is the size-biased distribution of mixing group sizes, and for $c \geq 2$, $\mathbb{P}(\tilde{W}^{(n)} = c - 1)$ is the probability that an event is a mixing event multiplied by the probability that a mixing event involving a given infective is of size c. It is then assumed that a mixing group of size c involving an infective produces $c - 1$ new infections, the maximum number of new infections that can be produced from a mixing group of size c.

The proof that $W^{(n)}(x^{(n)}, y^{(n)}) \overset{st}{\leq} \tilde{W}^{(n)}$ is as follows. Remember that $Q_c^{(n)}(l, i | x^{(n)}, y^{(n)})$, defined in (4.2), is the probability that a mixing group of size c in a population of size n containing $x^{(n)}$ susceptibles and $y^{(n)}$ infectives contains l susceptibles and i infectives. Note that for $v > 0$, the probability that such a mixing group results in v new infectives, $\pi_c(v; l, i)$, is not equal to 0 only if $i > 0$ and $v \leq l$. Given that there are $y^{(n)}$ infectives in the population, the probability that a mixing group of size c includes at least one infective is at most $cy^{(n)}/n$, and therefore, for $w = 0, 1, \ldots$,

$$
\mathbb{P}\left(W^{(n)}(x^{(n)}, y^{(n)}) \geq w\right)
$$

$$
= \frac{1}{\gamma + \beta^{(n)}} \sum_{v=w}^{n-1} q^{(n)}(x^{(n)}, y^{(n)}, v)
$$

$$
= \frac{1}{\gamma + \beta^{(n)}} \sum_{v=w}^{n-1} \frac{n\lambda}{y^{(n)}} \sum_{c=v+1}^{n} p_C^{(n)}(c) \sum_{i,j} Q_c^{(n)}(c - i - j, i | x^{(n)}, y^{(n)}) \pi_c(v; c - i - j, i)
$$

$$
= \frac{1}{\gamma + \beta^{(n)}} \sum_{c=w+1}^{n} p_C^{(n)}(c) \frac{n\lambda}{y^{(n)}} \sum_{i,j} Q_c^{(n)}(c - i - j, i | x^{(n)}, y^{(n)}) \left\{ \sum_{v=w}^{c-1} \pi_c(v; c - i - j, i) \right\}
$$

$$
\leq \frac{1}{\gamma + \beta^{(n)}} \sum_{c=w+1}^{n} p_C^{(n)}(c) \frac{n\lambda}{y^{(n)}} \times \frac{cy^{(n)}}{n}
$$

$$
= \frac{\lambda}{\gamma + \beta^{(n)}} \sum_{c=w+1}^{n} c p_C^{(n)}(c)
$$

$$
= \frac{\lambda \mu_C^{(n)}}{\gamma + \beta^{(n)}} \sum_{c=w+1}^{n} \mathbb{P}(\tilde{C}^{(n)} = c) = \mathbb{P}\left(\tilde{W}^{(n)} \geq w\right),
$$

as required.

Hence we can couple $\left\{W_k^{(n)}\left(A_{k-1}^{(n)}, G_{k-1}^{(n)}\right)\right\}$ and $\{\tilde{W}_k^{(n)}\}$ so that, for all $k = v_n + 1, v_n + 2, \ldots$,

$$
W_k^{(n)}\left(A_{k-1}^{(n)}, G_{k-1}^{(n)}\right) \leq \tilde{W}_k^{(n)}.
$$

Recall that $\mathbb{E}[(C^{(n)})^2] \to \mathbb{E}[C^2]$ as $n \to \infty$, where $\mathbb{E}[C^2] < \infty$. By the weak law of large numbers for triangular arrays, we have that

$$
\frac{1}{\lfloor \delta_1 n \rfloor - v_n} \sum_{k=v_n+1}^{\lfloor \delta_1 n \rfloor} |\tilde{W}_k^{(n)}| \overset{\mathrm{p}}{\longrightarrow} \frac{\gamma}{\gamma + \beta} + \frac{\beta}{\gamma + \beta} \frac{\mathbb{E}[C^2]}{\mu_C} \quad \text{as } n \to \infty, \tag{5.15}
$$

where the right-hand side of (5.15) is less than $\mathbb{E}[C^2]/\mu_C$. As noted at the start of the proof, $G_k^{(n)} = Y_k = L_k^{(n)}$ ($k = 1, 2, \ldots, v_n$) with probability tending to 1 as $n \to \infty$, where $Y_k = m + \sum_{j=1}^{k} \tilde{Z}_j$. Therefore, for any ϵ_2 satisfying $\delta_1 \mathbb{E}[C^2]/\mu_C < \epsilon_2 < \epsilon_1$, we have that

$$\mathbb{P}\left(A^{(n)}_{\lfloor\delta_1 n\rfloor}>(1-\epsilon_1)n\right)\geq\mathbb{P}\left(\left\{m+\sum_{j=1}^{v_n}|\tilde{Z}_j|\leq(\epsilon_1-\epsilon_2)n\right\}\cap\bigcap_{k=1}^{v_n}\left\{G^{(n)}_k=Y_k=L^{(n)}_k\right\}\right)$$

$$\times\mathbb{P}\left(\sum_{k=v_n+1}^{\lfloor\delta_1 n\rfloor}|\tilde{W}^{(n)}_k|\leq\epsilon_2 n\right)\to 1\quad\text{as }n\to\infty.$$

In Section 4.3, we showed that, during the early stages of the epidemic, the contribution to the spread of the disease from mixing events containing more than one non-susceptible individual is negligible, and whilst the number of susceptibles remains above $(1-\epsilon_1)n$ we can similarly bound the contribution from mixing events with multiple non-susceptible individuals. Following the proof of Lemma 4.1, we have for $w=1,2,\ldots,\lceil n/2\rceil$ that $q^{(n)}\big(x^{(n)},y^{(n)},w\big)\geq q^{(n)}_1\big(x^{(n)},y^{(n)},w\big)$, and using (4.12) and (4.13), for $x^{(n)}\geq(1-\epsilon_1)n$, we have that

$$q^{(n)}_1\big(x^{(n)},y^{(n)},w\big)\geq\frac{n\lambda}{y^{(n)}}\sum_{c=w+1}^{\lfloor n/2\rfloor}p^{(n)}_C(c)\frac{cy^{(n)}}{n}\left[1-\frac{2(c-1)\epsilon_1 n}{n}\right]\pi_c(w;c-1,1)$$

$$=\lambda\sum_{c=w+1}^{\lfloor n/2\rfloor}cp^{(n)}_C(c)\left[1-2(c-1)\epsilon_1\right]\pi_c(w;c-1,1)$$

$$\geq\beta^{(n)}\psi^{(n)}_w$$

for all sufficiently large n, as, for such n, $\psi^{(n)}_w=0$ for all $w>\lfloor n/2\rfloor$. Hence, for $k=v_n+1,v_n+2,\ldots,\lfloor\delta_1 n\rfloor$, provided that $A^{(n)}_{k-1}\geq(1-\epsilon_1)n$, we can couple $W^{(n)}_k\big(A^{(n)}_{k-1},G^{(n)}_{k-1}\big)$ and $\hat{Z}^{(n)}_k$, defined in (5.13) and (5.9) respectively, so that

$$\hat{Z}^{(n)}_k\leq W^{(n)}_k\big(A^{(n)}_{k-1},G^{(n)}_{k-1}\big)$$

for all sufficiently large n. Specifically, we couple deaths (downward steps) in $\mathcal{L}^{(n)}$ with removals in $\mathcal{G}^{(n)}$, so $W^{(n)}_k\big(A^{(n)}_{k-1},G^{(n)}_{k-1}\big)=-1$ if and only if $\hat{Z}^{(n)}_k=-1$. For $w=1,2,\ldots$, if $W^{(n)}_k\big(A^{(n)}_{k-1},G^{(n)}_{k-1}\big)=w$ then we set $\hat{Z}^{(n)}_k=w$ with probability $\psi^{(n)}_w/q^{(n)}\big(x^{(n)},y^{(n)},w\big)$ and set $\hat{Z}^{(n)}_k=0$ otherwise. Provided that $A^{(n)}_{\lfloor\delta_1 n\rfloor}\geq(1-\epsilon_1)n$, it then immediately follows by induction that for $k=v_n+1,v_n+2,\ldots,\lfloor\delta_1 n\rfloor$,

$$G^{(n)}_k=G^{(n)}_{k-1}+W^{(n)}_k\big(A^{(n)}_{k-1},G^{(n)}_{k-1}\big)\geq L^{(n)}_{k-1}+\hat{Z}^{(n)}_k=L^{(n)}_k,$$

and (5.14) holds.   □

The final step is to couple $\mathcal{S}^{(n)}$ to $\mathcal{G}^{(n)}$. By definition, the processes $\mathcal{S}^{(n)}$ and $\mathcal{G}^{(n)}$ coincide for the first v_n events, so $\big(X^{(n)}_{v_n},Y^{(n)}_{v_n}\big)=\big(A^{(n)}_{v_n},G^{(n)}_{v_n}\big)$. Remember, for $k=v_n+1,v_n+2,\ldots,\lfloor\delta_1 n\rfloor$, that $\kappa_n(k)$, defined in (5.4), denotes the number of events that have occurred in $\mathcal{S}^{(n)}$ up to and including the kth event in $\mathcal{G}^{(n)}$, with $\kappa_n(v_n)=v_n$ by definition. It is helpful to consider the epidemic processes $\mathcal{E}^{(n)}$ ($\mathcal{S}^{(n)}$) and $\mathcal{E}^{(n)}_G$ ($\mathcal{G}^{(n)}$) and to note that the rates at which events occur are $\left\{\gamma+\sum_{v=0}^{n-1}q^{(n)}(x^{(n)},y^{(n)},v)\right\}y^{(n)}$ and $\{\gamma+\beta^{(n)}\}y^{(n)}$, respectively. Therefore, the rate of occurrence of an event which results in $w=1,2,\ldots,n-1$ infections, when the population is in state $\big(x^{(n)},y^{(n)}\big)$, is $q^{(n)}\big(x^{(n)},y^{(n)},w\big)y^{(n)}$ in both $\mathcal{E}^{(n)}$ and $\mathcal{E}^{(n)}_G$. Similarly, the rate at

which a removal occurs in state $\left(x^{(n)}, y^{(n)}\right)$ is $\gamma y^{(n)}$. Thus, the only difference in event rates is for infection events which produce no infections where the rates are $q^{(n)}\left(x^{(n)}, y^{(n)}, 0\right)$ and $\beta^{(n)} - \sum_{v=1}^{n-1} q^{(n)}\left(x^{(n)}, y^{(n)}, v\right)$ in $\mathcal{E}^{(n)}$ and $\mathcal{E}_G^{(n)}$, respectively. Hence, if $W_k^{(n)}\left(A_{k-1}^{(n)}, G_{k-1}^{(n)}\right) \neq 0$, we set $\kappa_n(k) = \kappa_n(k-1) + 1$, and

$$
\begin{aligned}
X_{\kappa_n(k)}^{(n)} &= X_{\kappa_n(k-1)}^{(n)} - W_k^{(n)}\left(A_{k-1}^{(n)}, G_{k-1}^{(n)}\right) 1_{\left\{W_k^{(n)}\left(A_{k-1}^{(n)}, G_{k-1}^{(n)}\right)\right\}}, \\
Y_{\kappa_n(k)}^{(n)} &= Y_{\kappa_n(k-1)}^{(n)} + W_k^{(n)}\left(A_{k-1}^{(n)}, G_{k-1}^{(n)}\right).
\end{aligned}
\tag{5.16}
$$

That is, each event which leads to a change in the state of the population in $\mathcal{E}_G^{(n)}$ $\left(\mathcal{G}^{(n)}\right)$ has a corresponding event in $\mathcal{E}^{(n)}$ $\left(\mathcal{S}^{(n)}\right)$. Similarly, if $W_k^{(n)}\left(A_{k-1}^{(n)}, G_{k-1}^{(n)}\right) = 0$, we set $\kappa_n(k) = \kappa_n(k-1) + 1$ with probability $q^{(n)}\left(x^{(n)}, y^{(n)}, 0\right) / \left\{\beta^{(n)} - \sum_{v=1}^{n-1} q^{(n)}\left(x^{(n)}, y^{(n)}, v\right)\right\}$ and (5.16) holds with $\left(X_{\kappa_n(k)}^{(n)}, Y_{\kappa_n(k)}^{(n)}\right) = \left(X_{\kappa_n(k-1)}^{(n)}, Y_{\kappa_n(k-1)}^{(n)}\right)$; otherwise we set $\kappa_n(k) = \kappa_n(k-1)$, corresponding to no event in $\mathcal{E}^{(n)}$ $\left(\mathcal{S}^{(n)}\right)$ and a ghost event in $\mathcal{E}_G^{(n)}$ $\left(\mathcal{G}^{(n)}\right)$. Thus there exists $\kappa_n(\lfloor \delta_1 n \rfloor) \leq \lfloor \delta_1 n \rfloor$ such that

$$
Y_{\kappa_n(\lfloor \delta_1 n \rfloor)}^{(n)} = G_{\lfloor \delta_1 n \rfloor}^{(n)}. \tag{5.17}
$$

Proof of Theorem 3.2(b). Let $v_n = \lfloor \log n \rfloor$ and $u_n = \lfloor 3v_n / \mathbb{P}(Z \neq 0) \rfloor$, so (u_n) satisfies the conditions stated in Lemma 4.2. Let ϵ satisfy (5.6), and fix δ_1 and $\epsilon_1 < \epsilon$ such that $0 < \delta_1 < \epsilon_1 \mu_C / \mathbb{E}[C^2]$ (i.e. satisfying (5.7)). It follows from (5.3) that, for $R_0 > 1$, $\mathbb{E}[Z] > 0$ and for $T^{(n)} \geq v_n$ with probability tending to 1 as $n \to \infty$, $Y_{v_n} > v_n \mathbb{E}[Z]/2$.

Given $\delta > 0$ satisfying (5.10), it follows from (5.17) and Lemma 5.2 that, with probability tending to 1 as $n \to \infty$, for all $k = v_n + 1, v_n + 2, \ldots, \lfloor \delta_1 n \rfloor$,

$$
Y_{\kappa_n(k)}^{(n)} = G_k^{(n)} \geq L_k^{(n)}.
$$

By setting $b_n = v_n \mathbb{E}[Z]/2$ in Lemma 5.1, we have that, as $n \to \infty$,

$$
\mathbb{P}\left(Y_{\kappa_n(\lfloor \delta_1 n \rfloor)}^{(n)} \geq \lfloor \delta n \rfloor, \min_{v_n \leq k < \kappa_n(\lfloor \delta_1 n \rfloor)} \left\{Y_k^{(n)}\right\} > 0 \,\middle|\, T^{(n)} \geq v_n\right) \to 1. \tag{5.18}
$$

Given that $T^{(n)} \geq Y_{\kappa_n(\lfloor \delta_1 n \rfloor)}^{(n)}$, (3.14) follows immediately. $\qquad\square$

Finally, note that Corollary 3.2 follows immediately from (5.18) as $\sup_{t \geq 0} I^{(n)}(t) \geq Y_{\kappa_n(\lfloor \delta_1 n \rfloor)}^{(n)}$.

6. Concluding comments

As noted in the introduction, the aims of this paper are to provide a rigorous justification for the approximating branching process introduced in Ball and Neal [5] and a proof of a key result (Theorem 3.2 of this paper) required for a central limit theorem in [5] for the size of a major outbreak for epidemics with few initial infectives. The latter clearly requires a limit theorem. A limit theorem is also a common approach to rigorously justifying a branching process, but for practical purposes it is often useful to have information concerning the accuracy of the approximation for finite population size n, as given for example by a bound on the total variation distance between the epidemic $\mathcal{E}^{(n)}$ in a population of size n and the limiting branching

process $\mathcal{B}$. A detailed analysis of such accuracy of approximation is beyond the scope of the paper, so here we make a few very brief comments.

Recall from Section 4.5 that H_n denotes the first event at which a mismatch occurs between the embedded discrete jump processes of $\mathcal{E}^{(n)}$ and $\mathcal{B}$. It follows immediately from results in the proof of Lemma 4.2, using the notation in that lemma and its proof, that

$$\mathbb{P}(H_n \le u_n) \le \mathbb{P}(A_{n,0}^c) + \mathbb{P}(A_{n,1}^c) + \mathbb{P}(A_{n,2}^c), \tag{6.1}$$

thus yielding a bound on the total variation distance between $\mathcal{E}^{(n)}$ and $\mathcal{B}$ over the first u_n events for quantities that do not depend on the times of those events. The latter can be included by adding $\mathbb{P}(A_{n,3}^c)$ to the right-hand side of (6.1). Bounds for $\mathbb{P}(A_{n,i}^c)$ ($i = 0, 1, 2, 3$) can be obtained using results given in the proof of Lemma 4.2. For approximation purposes a source of inaccuracy can be removed by using a branching process defined analogously to $\mathcal{B}$ but with C replaced by $C^{(n)}$.

Acknowledgement

Funding information

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] BAILEY, N. T. J. (1975). *The Mathematical Theory of Infectious Diseases and Its Applications*, 2nd edn. Griffin, London.

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

[3] BALL, F. AND DONNELLY, P. (1995). Strong approximations for epidemic models. *Stoch. Process. Appl.* **55**, 1–21.

[4] BALL, F. AND NEAL, P. (2017). The asymptotic variance of the giant component of configuration model random graphs. *Ann. Appl. Prob.* **27**, 1057–1092.

[5] BALL, F. AND NEAL, P. (2022). An epidemic model with short-lived mixing groups. *J. Math. Biol.* **85**, article no. 63.

[6] BALL, F. AND O'NEILL, P. (1994). Strong convergence of stochastic epidemics. *Adv. Appl. Prob.* **26**, 629–655.

[7] BARTLETT, M. S. (1955). *An Introduction to Stochastic Processes*. Cambridge University Press.

[8] CORTEZ, R. (2022). *SIR model with social gatherings*. Preprint. Available at https://arxiv.org/abs/2203.08260.

[9] DURRETT, R. (2010). *Probability: Theory and Examples*. Cambridge University Press.

[10] ETHIER, S. N. AND KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence*. John Wiley, Hoboken, NJ.

[11] JAGERS, P. (1975). *Branching Processes with Biological Applications*. John Wiley, London.

[12] KENDALL, D. (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.

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

[14] NERMAN, O. (1981). On the Convergence of Supercritical General (C-M-J) Branching Processes. *Z. Wahrscheinlichkeitsth.* **57**, 365–395.

[15] NEWMAN, M. E. J. (2002). Spread of epidemic disease on networks. *Phys. Rev. E* **66**, article no. 016128.

[16] POLLETT, P. K. (1990). On a model for interference between searching insect parasites. *J. Austral. Math. Soc. B* **32**, 133–150.

[17] WHITTLE, P. (1955). The outcome of a stochastic epidemic—a note on Bailey's paper. *Biometrika* **42**, 116–122.

[18] WILLIAMS, T. (1971). An algebraic proof of the threshold theorem for the general stochastic epidemic. *Adv. Appl. Prob.* **3**, 223.