COUPLING OF TWO SIR EPIDEMIC MODELS WITH VARIABLE SUSCEPTIBILITIES AND INFECTIVITIES

PETER NEAL,* *University of Manchester*

Abstract

The *variable generalised stochastic epidemic* model, which allows for variability in both the susceptibilities and infectivities of individuals, is analysed. A very different epidemic model which exhibits variable susceptibility and infectivity is the *random-graph epidemic model*. A suitable coupling of the two epidemic models is derived which enables us to show that, whilst the epidemics are very different in appearance, they have the same asymptotic final size distribution. The coupling provides a novel approach to studying random-graph epidemic models.

Keywords: SIR epidemic; random graph; branching process; central limit theorem; coupling

2000 Mathematics Subject Classification: Primary 92D30
Secondary 60F05; 05C80

1. Introduction

The generalised stochastic epidemic (GSE) model is the most widely studied SIR (susceptible $\rightarrow$ infected $\rightarrow$ removed) epidemic model in a closed community. The GSE model assumes that infectives have independent and identically distributed infective periods distributed according to an arbitrary, but specified, nonnegative random variable Q. Whilst infective, individuals make infective contact at the points of an independent, homogeneous Poisson point process with rate $\lambda > 0$. Each infective contact is with an individual who is chosen uniformly at random from the population. Note that GSE usually refers to the general stochastic epidemic, which is the special case where $Q \sim \mathrm{Exp}(\gamma)$ for some $\gamma > 0$.

The GSE model assumes that there is variability in the infectivity of different individuals. The homogeneous mixing assumption for infective contacts implies that all individuals are equally susceptible to the disease. It is more realistic to assume that individuals have both variable susceptibilities and infectivities. Therefore, we consider the *variable generalised stochastic epidemic* (*VGSE*), which incorporates both of these elements. A number of authors have considered variable susceptibility and infectivity, in particular Ball [3] and Scalia-Tomba [19]. However, in these cases it was assumed that the susceptibility and infectivity of each individual are independent, whereas we allow explicitly for dependence between susceptibility and infectivity. Another model which allows for variability in both susceptibility and infectivity of individuals, but in rather specialised ways, is the two-stage epidemic model of [4].

The VGSE model is closely related to the multitype SIR epidemic model; see, for example, [5], [6], and [17]. In fact, the VGSE model is the continuous limit of an important special

case of the multitype model of [17, Section 6], where the rate at which an infective makes infective contacts with susceptibles of type j only depends upon the susceptibility of type-j individuals. The model in [6] is the above model with susceptibility independent of type. The VGSE model is also related to a vaccine response model introduced in [10]. In [10], a proportion of the individuals in a homogeneous population are given a vaccine which affects an individual's susceptibility and infectivity. The emphasis of [10] is on the critical vaccination coverage for the population. More recently, in [14], the exact final size distribution of epidemics with variable susceptibilities and infectivities has been derived, the final size of an epidemic being defined as the total number of initially susceptible individuals who are ultimately infected during the course of the epidemic. Therefore, we focus on the asymptotic final size distribution as the initial number of susceptibles, n, goes to ∞.

Very different epidemic models which exhibit variability in the susceptibilities and infectivities of individuals are epidemics upon random graphs. These models are particularly relevant to social and computer networks, where an individual's susceptibility and infectivity are linked to his activity levels and/or connectedness. Prime examples are sexually transmitted diseases and e-mail viruses. In such models there is assumed to be an underlying (random) graph connecting the individuals in the population. In particular, individuals are associated with vertices in the graph and two individuals are said to be acquaintances (connected) if and only if an edge exists between the corresponding two vertices in the graph. Individuals, whilst infective, can only make infective contact with their acquaintances. For the Bernoulli random graph an edge exists between two vertices with probability α, $0 < \alpha \leq 1$, independent of the remainder of the graph. We shall consider epidemics upon variable Bernoulli random graphs, that is, ones in which an edge exists between two vertices with a probability dependent upon the connectivity of the vertices in question but independent of the remainder of the graph. The connectivity random variables for the vertices are assumed to be independent and identically distributed. A fuller description of the model will be given in Section 5.

A coupling of the two epidemic processes is constructed. In particular, we show that, for any variable Bernoulli random-graph epidemic, there exists a VGSE with the same asymptotic final size distribution. Therefore, it is sufficient to analyse the VGSE in detail and translate the results in terms of the random-graph epidemic. The fact that we can couple these two rather different epidemic models may at first seem rather surprising. However, the infection processes in the models are similar in nature. In both cases, the probability that, whilst infective, an infective makes an infective contact with a susceptible is dependent only upon the infectivity (and connectivity) of the (potential) infector and the susceptibility/connectivity of the susceptible. For the VGSE model this consists of just the infection probability between the two individuals. For the random-graph epidemic model this consists of two parts: first the probability that the individuals are acquainted/connected, and second, given they are acquainted, the probability that the infective infects the susceptible. Furthermore, in constructing the random-graph epidemic we combine these two steps and the link between the two epidemics appears more natural.

The paper is structured as follows. A description of the VGSE model is given in Section 2. In Section 3 a branching process approximation for minor epidemic outbreaks is obtained for the VGSE. In Section 4 major epidemic outbreaks are considered and a central limit theorem for the number infected in a major epidemic outbreak derived. In Section 5 the random-graph epidemic model and the coupling of the two epidemic processes are described. The asymptotic results of Sections 3 and 4 can then be applied to the random-graph epidemic model. Finally, a brief discussion of extensions of the current work is presented in Section 6.

2. Model description

Consider a closed population consisting of a initial infectives and n initial susceptibles. Label the initial infectives individuals $-(a-1), -(a-2), \ldots, 0$, and label the initial susceptibles individuals $1, 2, \ldots, n$. Suppose that an initial infective i, $-(a-1) \leq i \leq 0$, has infective period $\tilde{Q}_i \overset{\mathrm{D}}{=} \tilde{Q}$, for some arbitrary, but specified, nonnegative random variable $\tilde{Q}$ (where '$\overset{\mathrm{D}}{=}$' denotes equality in distribution). Let $(D_1, Q_1), (D_2, Q_2), \ldots, (D_n, Q_n)$ be independent and identically distributed according to (D, Q), where D and Q are arbitrary, but specified, nonnegative random variables. We assume that D and Q are dependent and we will make assumptions upon the distributions of D and Q as required. For $1 \leq i \leq n$, let D_i and Q_i respectively denote the susceptibility and the infective period of individual i.

For $-(a-1) \leq i \leq n$ and $1 \leq j \leq n$, we assume that, during his infective period, individual i makes infective contact with individual j at the points of a homogeneous Poisson point process with rate D_j/n. A susceptible individual is infected when he is contacted by an infective contact. Otherwise the infective contact has no effect. Therefore, whilst infective, individual i has probability $1 - \exp(-Q_i D_j/n)$ of making at least one infective contact with individual j.

The epidemic can be constructed in real time as follows. We assume that the initial infectives become infective at time 0. An individual i, say, who becomes infective at time t, say, is infective for the time period $(t, t + Q_i]$ (or $(0, \tilde{Q}_i]$ if individual i is an initial infective). During its infective period, individual i makes infective contacts at the points of a homogeneous Poisson point process with rate $\lambda_n = (1/n) \sum_{i=1}^{n} D_i$. For $1 \leq j \leq n$, the probability that individual j is contacted by a given infective contact is $D_j/n\lambda_n$. There is the possibility that some individuals are totally immune to the disease (prior to infection), corresponding to $D = 0$. Hence, we assume that $\mathrm{P}(D = 0) \geq 0$ and that $\mathrm{E}[D] > 0$. Therefore, it is possible that $\lambda_n = 0$ with all the initial susceptibles being immune, in which case no infection can take place. For $\lambda_n = 0$, we employ the convention that $D_j/n\lambda_n = 1/n$, $1 \leq j \leq n$. The following observation will prove useful throughout. For all $m \in \mathbb{N}$, if $\mathrm{E}[D^m] < \infty$ then

$$\mathrm{E}\left[\left(\frac{D_1}{\lambda_n}\right)^m\right] \to \frac{\mathrm{E}[D^m]}{\mathrm{E}[D]^m} \quad \text{as } n \to \infty. \tag{2.1}$$

In the special case $D \equiv \lambda$, for some $\lambda > 0$, the model is the generalised stochastic epidemic model.

The Sellke construction of the epidemic based upon [20] and [18] is an alternative approach but is equivalent in terms of the final size distribution of the epidemic process. We apply the Sellke construction to the VGSE as follows. Rather than considering infectives making infective contacts, we consider the susceptibles. We assume that each susceptible has a threshold for infection. That is, an individual accumulates exposure to infection and is susceptible whilst the accumulated infection is below the infective threshold. However, once the total amount of accumulated infection exceeds the threshold, the susceptible becomes infected. This description of the epidemic process is most easily (in terms of analysis) made fully rigorous by using an embedding approach introduced in [18].

For $1 \leq i \leq n$, let $H_i \sim \mathrm{Exp}(D_i)$, where H_i denotes the infective threshold of individual i. For $1 \leq i \leq n$ and $t \geq 0$, let $\chi_i(t) = 1_{\{H_i \leq t\}}$. Let $R_n(t) = \sum_{i=1}^{n} \chi_i(t)$ and $A_n(t) = \sum_{i=1}^{n} \chi_i(t) Q_i$. Thus, $R_n(t)$ and $A_n(t)$ denote the total number of initial susceptibles infected and the corresponding sum of infective periods when each individual is exposed to t units of global infective pressure. Let $S_0^n = n \bar{S}_0^n = \sum_{i=-(a-1)}^{0} \tilde{Q}_i$ denote the sum of the infective periods of the initial infectives.

The epidemic process can be constructed as follows. Let the initial infectives form generation 0 and, for $k \geq 1$, let generation k comprise those individuals infected by the infectives in generation $k-1$. For $k \geq 1$, let $T_k^n = n\bar{T}_k^n$ and $S_k^n = n\bar{S}_k^n$ respectively denote the total number of initial susceptibles infected in the first k generations of the epidemic and the corresponding sum of these individuals' infective periods, with $T_0^n = 0$. In particular, for $1 \leq i \leq n$ and $k \geq 0$, an individual i, say, is said to be infected by the individuals in generation k, and hence be an infective in generation $k+1$, if $\bar{S}_{k-1}^n < H_i \leq \bar{S}_k^n$ with the convention $S_{-1}^n = 0$. Thus, for $k \geq 1$,

$$S_k^n = S_0^n + A_n(\bar{S}_{k-1}^n),$$
$$T_k^n = R_n(\bar{S}_{k-1}^n).$$

Consequently, the above sequence stops at generation k^*, where $k^* = \min\{k : (T_{k+1}^n, S_{k+1}^n) = (T_k^n, S_k^n)\}$. Since n is finite, k^* is well defined. Let $T_\infty^n = n\bar{T}_\infty^n$ and $S_\infty^n = n\bar{S}_\infty^n$ respectively denote the final size of the epidemic (the total number of initial susceptibles infected during the course of the epidemic) and the final severity of the epidemic (the total sum of the infective periods of all individuals infected in the epidemic). Then $\bar{S}_\infty^n$ and $\bar{T}_\infty^n$ respectively satisfy

$$\bar{S}_\infty^n = \bar{S}_0^n + \frac{1}{n}A_n(\bar{S}_\infty^n),$$
$$\bar{T}_\infty^n = \frac{1}{n}R_n(\bar{S}_\infty^n).$$

In particular,

$$\bar{S}_\infty^n = \min\left\{s \geq 0 : s = \bar{S}_0^n + \frac{1}{n}A_n(s)\right\}. \tag{2.2}$$

Finally, since the results of this paper are asymptotic we shall consider a sequence of epidemics, $\{E_n\}$, each indexed by the total number of initial susceptibles, n, as $n \to \infty$. For a fixed $n \geq 1$, the epidemic E_n is constructed as outlined above. Note that $\tilde{Q}$ and (D, Q) are assumed to be independent of n. For each n, we denote the total number of initial infectives by a^n. We shall consider the respective cases in which $a^n = a$ for $n \geq 1$ and $a^n \to \infty$ as $n \to \infty$. For ease of notation, the index n will usually be omitted.

3. Branching process approximation

For a fixed $n \geq 1$, we shall construct the epidemic E_n and a suitable approximating branching process $\mathcal{B}_n$ on a common probability space $(\Omega, \mathcal{F}, P)$. We shall then define the branching process $\mathcal{B}$ as the limit of the sequence of branching processes $\mathcal{B}_n$ as $n \to \infty$. Letting Z denote the total progeny (excluding the initial ancestors) of the branching process $\mathcal{B}$, we shall show in Theorem 3.1 that $T_\infty^n \overset{D}{\to} Z$ as $n \to \infty$ (where '$\overset{D}{\to}$' denotes convergence in distribution). In Corollary 3.1 we summarise the useful results which can be obtained from the coupling of the epidemics E_n to the branching process $\mathcal{B}$. We shall require that there exist a $\kappa > 0$ such that $E[D^{1+\kappa}] < \infty$, but impose no moment conditions upon Q. Throughout this section it is assumed that $a^n = a$, $n \geq 1$. (For the case in which $a^n \to \infty$ as $n \to \infty$, the results of Section 4 are sufficient.)

Fix $n \geq 1$ and construct $\mathcal{B}_n$ as follows. We assume that there are a initial ancestors, born at time $t = 0$. Label the individuals $-(a-1), -(a-2), \ldots, 0$, and let the ith initial ancestor have lifetime $V_i^n = \tilde{Q}_i$, $-(a-1) \leq i \leq 0$. Let $D_1, D_2, \ldots, D_n$ be independent and identically

distributed according to D, with $\boldsymbol{D}^n = (D_1, D_2, \ldots, D_n)$. For $1 \le i \le n$, let $P_i^n = D_i/n\lambda_n$. Let $\psi_1^n, \psi_2^n, \ldots$ be independent and identically distributed according to ψ^n, where

$$P(\psi^n = i \mid \boldsymbol{D}^n) = \begin{cases} P_i^n, & 1 \le i \le n, \\ 0, & \text{otherwise.} \end{cases}$$

For $i \ge 1$, let $G_i^n = D_{\psi_i^n}$ and let $V_i^n \stackrel{\mathrm{D}}{=} \hat{Q}^{G_i^n}$, where $\hat{Q}^x \stackrel{\mathrm{D}}{=} [Q_1 \mid D_1 = x]$ for $x \ge 0$. Let V_i^n denote the length of the lifetime of the ith individual born in $\mathcal{B}_n$. For $i = -(a-1), -(a-2), \ldots,$ let η_i^n be a homogeneous Poisson point process with rate λ_n. Suppose that the ith birth occurs at time t_i^n. Then the ith individual dies at time $t_i^n + V_i^n$, and whilst alive reproduces at the points of η_i^n.

Let $M_n = \min\{k > 1 : \psi_k^n \in \{\psi_1^n, \psi_2^n, \ldots, \psi_{k-1}^n\}\}$. The epidemic E_n can then be coupled to the branching process $\mathcal{B}_n$ up to time t_{M_n} as follows. Label the initial infectives individuals $-(a-1), -(a-2), \ldots, 0$, with respective infective periods $\tilde{Q}_{-(a-1)}, \tilde{Q}_{-(a-2)}, \ldots, \tilde{Q}_0$. Whilst infective, individual i, $-(a-1) \le i \le 0$, makes infective contacts at the points of η_i^n. For $1 \le i \le n$, assign to individual i susceptibility D_i. For $1 \le i < M_n$, let the individual contacted by the ith global infective contact be individual ψ_i^n. Set $Q_{\psi_i^n} = V_i^n \stackrel{\mathrm{D}}{=} \hat{Q}^{D_{\psi_i^n}}$. Whilst infective, individual ψ_i^n, $1 \le i < M_n$, makes infective contacts at the points of η_i^n. Therefore, the infective period and the infection rate of the ith infected individual, ψ_i^n, have the correct distribution. Furthermore, up until time t_{M_n} the infective contacts in the epidemic process and the births in the branching process are in one-to-one correspondence.

Let Z_n denote the total progeny of the branching process $\mathcal{B}_n$. Then $T_\infty^n = k$ for any k, $1 \le k < M_n$, if and only if $Z_n = k$. Therefore, we consider the random variable M_n before studying the limiting distribution of Z_n. The limiting distribution of M_n has been studied in detail in [13, Theorem 5.2], under extra conditions upon D. However, Lemma 3.1 is sufficient for our needs and is valid for more general choices of D.

Lemma 3.1. *Suppose that there exists a $\kappa > 0$ such that $\mathrm{E}[D^{1+\kappa}] < \infty$. Then, for any α with $0 < \alpha < \kappa/2(1+\kappa)$,*

$$P(M_n > n^\alpha) \to 1 \quad \text{as } n \to \infty.$$

Proof. Fix $\alpha < (1-\gamma)/2 < \kappa/2(1+\kappa)$. For $n \ge 1$, let

$$S_n = \bigcap_{i=1}^n \{D_i \le n^\gamma\} \cap \left\{\lambda_n \ge \frac{1}{2}\mathrm{E}[D]\right\}.$$

It is trivial using Markov's inequality to show that $P(S_n) \to 1$ as $n \to \infty$. For $2 \le k \le \lfloor n^\alpha \rfloor + 1$, it straightforwardly follows by induction that

$$P(M_n > k \mid S_n) \ge \prod_{i=1}^{k-1}\left(1 - \frac{in^\gamma}{n\,\mathrm{E}[D]/2}\right)$$

(where, for $x \in \mathbb{R}$, $\lfloor x \rfloor$ denotes the integer part of x). Therefore, $P(M_n > n^\alpha \mid S_n) \to 1$ as $n \to \infty$, and the lemma is proved.

Returning to $\mathcal{B}_n$, note that the offspring distributions of an initial ancestor and an individual born in the branching process are $\tilde{R}^n \sim \mathrm{Po}(\lambda_n \tilde{Q})$ and $R^n \sim \mathrm{Po}(\lambda_n V^n)$, respectively, where $V^n \stackrel{\mathrm{D}}{=} V_1^n$.

Let $V \stackrel{D}{=} \hat{Q}^G$, where G is a random variable with probability density function

$$f_G(x) = \frac{x}{E[D]} f_D(x) \quad \text{as } n \to \infty$$

and $E[V] = E[DQ]/E[D]$. Let $\mathcal{B}$ denote a branching process in which there are a initial ancestors and the offspring distributions of an initial ancestor and an individual born in the branching process are $\tilde{R} \sim \text{Po}(E[D]\tilde{Q})$ and $R \sim \text{Po}(E[D]V)$, respectively. It is a trivial extension of [15, Lemma 3.6] to show that if $\tilde{R}^n \stackrel{D}{\to} \tilde{R}$ and $R^n \stackrel{D}{\to} R$, then $Z_n \stackrel{D}{\to} Z$ as $n \to \infty$. Now suppose that $V^n \stackrel{D}{\to} V$ as $n \to \infty$. Then, since $\lambda_n \stackrel{P}{\to} E[D]$ (where '$\stackrel{P}{\to}$' denotes convergence in probability), $\lambda_n V^n \stackrel{D}{\to} E[D]V$ as $n \to \infty$. Therefore, by [15, Lemma 5.8], $\tilde{R}^n \stackrel{D}{\to} \tilde{R}$ and $R^n \stackrel{D}{\to} R$ as $n \to \infty$. Therefore, we proceed by firstly showing that $V^n \stackrel{D}{\to} V$ as $n \to \infty$, in Lemma 3.2, and secondly showing that $T_\infty^n \stackrel{D}{\to} Z$ as $n \to \infty$, in Theorem 3.1.

Lemma 3.2. $V^n \stackrel{D}{\to} V$ as $n \to \infty$.

Proof. For $x, y \geq 0$, let $H(x; y) = P(Q_1 \leq x \mid D_1 = y)$. Then, for all $n \geq 1$ and $y \geq 0$,

$$P(V^n \leq y) = E_{D^n}[P(V^n \leq y \mid D^n)] = E\left[\frac{1}{n} \sum_{i=1}^n H(y; D_i) \frac{D_i}{\lambda_n}\right] = E\left[H(y; D_1) \frac{D_1}{\lambda_n}\right]. \quad (3.1)$$

By (2.1), the right-hand side of (3.1) converges, as $n \to \infty$, to

$$E\left[H(y; D_1) \frac{D_1}{E[D]}\right] = P(V \leq y),$$

and the lemma is proved. $\qquad \blacksquare$

Theorem 3.1. $T_\infty^n \stackrel{D}{\to} Z$ as $n \to \infty$.

Proof. An immediate consequence of Lemma 3.2 is that $Z_n \stackrel{D}{\to} Z$ as $n \to \infty$ (see the discussion proceeding Lemma 3.2). Therefore, for all $k \in \mathbb{Z}$, by Lemma 3.1,

$$\begin{aligned}
P(T_\infty^n \leq k) &= P(T_\infty^n \leq k \mid M_n > k) P(M_n > k) + P(T_\infty^n \leq k \mid M_n \leq k) P(M_n \leq k) \\
&= P(Z_n \leq k \mid M_n > k) P(M_n > k) + P(T_\infty^n \leq k \mid M_n \leq k) P(M_n \leq k) \\
&\to P(Z \leq k) \quad \text{as } n \to \infty,
\end{aligned}$$

as required. $\qquad \blacksquare$

We say that a global epidemic occurs if, in the limit as $n \to \infty$, the epidemic process infects infinitely many susceptibles. Then, utilising standard branching process theory, we have the following immediate results concerning T_∞^n.

Corollary 3.1. *Let*

$$f(s) = E[s^R] = E[\exp((s-1) E[D]\hat{Q}^G)],$$
$$\tilde{f}(s) = E[s^{\tilde{R}}] = E[\exp((s-1) E[D]\tilde{Q})].$$

Then, as $n \to \infty$,

(a) *a global epidemic occurs with nonzero probability if and only if $E[R] = E[DQ] > 1$;*

(b) *the probability of a global epidemic is $1 - p_{\text{EXT}} = 1 - \tilde{f}(p)^a$, where p is the smallest solution to $f(s) = s$ in $[0, 1]$ and p_{EXT} is the extinction probability of the branching process $\mathcal{B}$; and*

(c) *the probability generating function of $T_\infty \overset{\text{D}}{=} Z$, the limiting final size of the epidemic, is $(\tilde{f}(h(s))/s)^a$, where $h(s)$ satisfies $h(s) = sf(h(s))$.*

4. Central limit theorem

The aim of this section is to establish a central limit theorem for the final proportion of the population infected by the epidemic in the event of a major epidemic outbreak. In particular, for $n \geq 1$ we shall say that a major epidemic occurs if $T_\infty^n > \log n$. Thus, $G_n = \{T_\infty^n > \log n\}$ denotes the event that a major epidemic occurs. We shall require that $\text{E}[D^2], \text{E}[Q^2] < \infty$.

The first step is to show that $[(\bar{T}_\infty^n, \bar{S}_\infty^n) \mid G_n] \overset{\text{P}}{\to} (\tau, \sigma)$ as $n \to \infty$, for some suitably defined $\tau, \sigma > 0$. The second step is to show that $[\sqrt{n}(\bar{T}_\infty^n - \tau) \mid G_n]$ converges in distribution to a Gaussian random variable as $n \to \infty$. Thus, we begin by deriving σ and τ and a suitable Gaussian random variable. For $s \geq 0$, let

$$r(s) = \text{E}[\chi_1(s)] = \text{E}[1 - \exp(-sD)],$$
$$a(s) = \text{E}[\chi_1(s)Q_1] = \text{E}[(1 - \exp(-sD))Q].$$

Therefore, $r(s)$ and $a(s)$ are respectively the mean probability that an individual is infected when each member of the population is exposed to s units of global infective pressure and the corresponding mean infective contribution. Furthermore, for all $n \geq 1$ and $s \geq 0$, $\text{E}[R_n(s)] = nr(s)$ and $\text{E}[A_n(s)] = na(s)$. We summarise some simple, but useful, results concerning $r(\cdot)$ and $a(\cdot)$ in the following lemma.

Lemma 4.1. *The functions $r(\cdot)$ and $a(\cdot)$ are increasing, concave, and continuously differentiable on $[0, \infty)$. Moreover, for $s \geq 0$, $r'(s) = \text{E}[De^{-sD}]$ and $a'(s) = \text{E}[QDe^{-sD}]$.*

For $\zeta \geq 0$, let

$$\sigma = \zeta \, \text{E}[\tilde{Q}] + a(\sigma). \tag{4.1}$$

Since $a(\cdot)$ is a nondecreasing, concave function with $a(\infty) = \text{E}[Q]$, for $\zeta > 0$ there exists a unique solution, $\sigma_\zeta \in (0, \infty)$, to (4.1). Moreover, $a'(\sigma_\zeta) < 1$. For $\zeta = 0$, $\sigma = 0$ is a solution to (4.1) and a second solution, $\sigma_0 > 0$, exists if and only if $a'(0) = \text{E}[QD] > 1$. Note that $a'(0) = R_0$; thus, by Corollary 3.1(a), this second solution exists if and only if the probability of a global epidemic is nonzero. Therefore, for $\zeta = 0$ and $a'(0) \leq 1$, the branching process approximation of Section 3 is sufficient for the analysis of T_∞^n. We shall thus assume that σ_0 exists. Finally, for $\zeta \geq 0$, let $\tau_\zeta = r(\sigma_\zeta)$.

For $n \geq 1$, $s \geq 0$, and $\alpha, \beta \in \mathbb{R}$, let

$$Z_n^{\alpha,\beta}(s) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{\chi_i^n(s)(\alpha + \beta Q_i) - (\alpha r(s) + \beta a(s))\}$$

$$= \alpha \frac{1}{\sqrt{n}}(R_n(s) - nr(s)) + \beta \frac{1}{\sqrt{n}}(A_n(s) - na(s)).$$

From the central limit theorem, we have the following trivial, but useful, result.

Lemma 4.2. *For all $s \geq 0$ and $\alpha, \beta \in \mathbb{R}$,*

$$Z_n^{\alpha,\beta}(s) \xrightarrow{\mathrm{D}} Z^{\alpha,\beta}(s) \quad \text{as } n \to \infty,$$

where $Z^{\alpha,\beta}(s)$ is a Gaussian random variable with mean 0 and variance

$$\mathrm{var}(\chi_1(s)(\alpha + \beta Q_1)) = \alpha^2 \, \mathrm{var}(\chi_1(s)) + 2\alpha\beta \, \mathrm{cov}(\chi_1(s), \chi_1(s)Q_1) + \beta^2 \, \mathrm{var}(\chi_1(s)Q_1).$$

Returning to τ_ζ and σ_ζ, we have the following result.

Lemma 4.3. *Suppose that $a/n \xrightarrow{\mathrm{P}} \zeta \geq 0$. Then, for $\zeta > 0$,*

$$(\bar{T}_\infty^n, \bar{S}_\infty^n) \xrightarrow{\mathrm{P}} (\tau_\zeta, \sigma_\zeta) \quad \text{as } n \to \infty$$

and, for $\zeta = 0$,

$$\min\{|(\bar{T}_\infty^n, \bar{S}_\infty^n)| \, |(\bar{T}_\infty^n, \bar{S}_\infty^n) - (\tau_\zeta, \sigma_\zeta)|\} \xrightarrow{\mathrm{P}} 0 \quad \text{as } n \to \infty,$$

where, for $x_1, x_2, y_1, y_2 \in \mathbb{R}$, $|(x_1, x_2) - (y_1, y_2)| = \sum_{i=1}^{2} |x_i - y_i|$.

Proof. It follows from Lemma 4.2 that, for all $t \geq 0$,

$$\frac{1}{n} A_n(t) \xrightarrow{\mathrm{P}} a(t) \quad \text{as } n \to \infty. \tag{4.2}$$

Then, since $A_n(t)$ and $a(t)$ are increasing, continuous functions of t, with $a(\infty) = \mathrm{E}[Q] < \infty$, it follows from (4.2) that

$$\sup_{t \geq 0} \left| \frac{1}{n} A_n(t) - a(t) \right| \xrightarrow{\mathrm{P}} 0 \quad \text{as } n \to \infty.$$

(See the proof of [7, Lemma 5.1]; see also [8], [4], and [17]). Similarly,

$$\sup_{t \geq 0} |n^{-1} R_n(t) - r(t)| \xrightarrow{\mathrm{P}} 0$$

as $n \to \infty$, and the lemma follows since $a/n \xrightarrow{\mathrm{P}} \zeta \geq 0$ as $n \to \infty$.

Therefore, for $\zeta > 0$, $[(\bar{T}_\infty^n, \bar{S}_\infty^n) \mid G_n] \xrightarrow{\mathrm{P}} (\tau_\zeta, \sigma_\zeta)$ and $\mathrm{P}(G_n) \to 1$ as $n \to \infty$. Considerably more work is required for the case in which $\zeta = 0$. We shall utilise a lower-bound branching process for the epidemic process. This idea originated in [21], and it is usually trivial to construct a suitable lower-bound branching process. However, the dependence between the susceptibility and infectivity of each individual causes added complications. The required result is proved in Corollary 4.2, via a series of intermediary lemmas.

Fix $n \geq 1$. For $1 \leq i \leq n$, let $\tilde{\psi}_i^n$ be drawn without replacement from $\{1, 2, \ldots, n\}$ as follows. For $1 \leq j \leq n$, let $\Psi_j^n = \{\tilde{\psi}_1^n, \tilde{\psi}_2^n, \ldots, \tilde{\psi}_{j-1}^n\}$. Then, for $1 \leq j, k \leq n$,

$$\mathrm{P}(\tilde{\psi}_j^n = k \mid \Psi_j^n) = \begin{cases} \dfrac{D_k}{\sum_{i \notin \Psi_j^n} D_i}, & k \notin \Psi_j^n, \\ 0, & k \in \Psi_j^n. \end{cases}$$

Therefore, for $1 \leq i \leq n$, let the ith susceptible infected in the epidemic E_n be the individual $\tilde{\psi}_i^n$ (assuming that there are at least i infections in E_n).

Lemma 4.4. *For all $n \geq 1$ and $1 \leq i \leq n-1$, $D_{\tilde{\psi}_{i+1}^n} \leq_{\mathrm{st}} D_{\tilde{\psi}_i^n}$, where '$\leq_{\mathrm{st}}$' means 'stochastically less than'.*

Proof. Fix $n \geq 1$. It is sufficient to show that $D_{\tilde{\psi}_2^n} \leq_{\mathrm{st}} D_{\tilde{\psi}_1^n}$, since the general result then follows straightforwardly by induction.

Let $\{\check{D}_i^n\}$ denote the set of $\boldsymbol{D}^n$ ordered such that

$$\check{D}_1^n \leq \check{D}_2^n \leq \cdots \leq \check{D}_n^n.$$

Let $\check{\psi}_i^n$ be drawn without replacement from $\{1, 2, \ldots, n\}$ with the probability that $\check{\psi}_i^n = j$ proportional to $\check{D}_j^n$ conditional upon $\check{\psi}_k^n \neq j$, $1 \leq k \leq i-1$. To prove the lemma, it is then sufficient to show that $\check{\psi}_2^n \leq_{\mathrm{st}} \check{\psi}_1^n$ or, equivalently, that $[\check{\psi}_2^n \mid \boldsymbol{\check{D}}^n] \leq_{\mathrm{st}} [\check{\psi}_1^n \mid \boldsymbol{\check{D}}^n]$ for all $\boldsymbol{\check{D}}^n$.

Note that, for all k, $1 \leq k \leq n$,

$$\mathrm{P}(\check{\psi}_1^n \leq k \mid \boldsymbol{\check{D}}^n) = \sum_{i=1}^{k} \frac{\check{D}_i^n}{n\lambda_n} \tag{4.3}$$

and

$$\mathrm{P}(\check{\psi}_2^n \leq k \mid \boldsymbol{\check{D}}^n) = \sum_{i=1}^{k} \sum_{j \neq i} \frac{\check{D}_i^n}{n\lambda_n - \check{D}_j^n} \frac{\check{D}_j^n}{n\lambda_n} = \sum_{i=1}^{k} \frac{\check{D}_i^n}{n\lambda_n} \sum_{j \neq i} \frac{\check{D}_j^n}{n\lambda_n - \check{D}_j^n}. \tag{4.4}$$

The lemma follows from (4.3) and (4.4) since $\sum_{j \neq i} \check{D}_j^n / (n\lambda_n - \check{D}_j^n)$ is nonincreasing in i. $\qquad \blacksquare$

Corollary 4.1. *For any sequence of positive integers b_n such that $b_n/n \to 0$ as $n \to \infty$, let $\delta_n = \sqrt{b_n/n}$. Then $\delta_n \to 0$ as $n \to \infty$ and*

$$\mathrm{P}\left(\frac{1}{n\lambda_n} \sum_{i \in \Psi_{b_n}^n} D_i \geq \delta_n \right) \to 0 \quad \text{as } n \to \infty.$$

Proof. Since $D_{\tilde{\psi}_1^n} \overset{\mathrm{D}}{=} D_{\psi_1^n}$, by Markov's inequality and Lemma 4.4,

$$\mathrm{P}\left(\frac{1}{n\lambda_n} \sum_{i \in \Psi_{b_n}^n} D_i \geq \delta_n \right) \leq \frac{1}{\delta_n} \mathrm{E}_{\boldsymbol{D}^n}\left[\sum_{i \in \Psi_{b_n}} \frac{D_{\tilde{\psi}_i^n}}{n\lambda_n} \right]$$

$$\leq \frac{b_n}{n\delta_n} \mathrm{E}\left[\frac{D_{\tilde{\psi}_1^n}}{\lambda_n} \right]$$

$$= \frac{b_n}{n\delta_n} \mathrm{E}\left[\sum_{i=1}^{n} \frac{D_i}{\lambda_n} \frac{D_i}{n\lambda_n} \right]. \tag{4.5}$$

By (2.1), the right-hand side of (4.5) converges to 0 as $n \to \infty$, and the corollary is proved. $\qquad \blacksquare$

Lemma 4.5. *For any sequence of positive integers b_n such that $b_n/n \to 0$ as $n \to \infty$,*

$$D_{\tilde{\psi}_{b_n}^n} \overset{\mathrm{D}}{\to} G \quad \text{as } n \to \infty.$$

Proof. For $n \geq 1$ and $1 \leq i \leq n$, couple $\tilde{\psi}_i^n$ and ψ_i^n as follows. Draw ψ_i^n from $\{1, 2, \ldots, n\}$ with probability $\mathrm{P}(\psi_i^n = j) = D_j/(n\lambda_n)$. If $\psi_i^n \notin \Psi_i^n$ then set $\tilde{\psi}_i^n = \psi_i^n$; otherwise, draw $\tilde{\psi}_i^n$ from $\{1, 2, \ldots, n\} \setminus \Psi_i^n$ with probability $\mathrm{P}(\tilde{\psi}_i^n = j) = D_j / \sum_{k \notin \Psi_i^n} D_k$. Following the proof of Lemma 3.2, $D_{\psi_1^n} = G^n \overset{\mathrm{D}}{\to} G$ as $n \to \infty$. The lemma then follows by

Corollary 4.1 and [11, Theorem 4.1], since

$$P(\tilde{\psi}_{b_n}^n \neq \psi_{b_n}^n) = E[P(\tilde{\psi}_{b_n}^n \neq \psi_{b_n}^n \mid \boldsymbol{D}^n)]$$

$$= E\left[\sum_{i \in \Psi_{b_n}^n} \frac{D_{\tilde{\psi}_i^n}}{n\lambda_n} \right] \to 0 \quad \text{as } n \to \infty.$$

We are almost in a position to show that $[\bar{T}_\infty^n \mid G_n] \xrightarrow{P} \tau_0$ as $n \to \infty$. However, we shall require the following assumption. For any sequence of positive integers b_n such that $b_n/n \to 0$ as $n \to \infty$, there exists a sequence of random variables $\check{Q}^{n,b_n}$ with the following properties. Firstly, $\check{Q}^{n,b_n} \xrightarrow{D} \hat{Q}^G \equiv V$ as $n \to \infty$ and, secondly, $\check{Q}^{n,b_n} \leq_{\text{st}} \hat{Q}^{D_{\tilde{\psi}_i^n}}$ for all i, $1 \leq i \leq b_n$. Note that if, for all $x < y$, $\hat{Q}^x \leq_{\text{st}} \hat{Q}^y$ or $\hat{Q}^y \leq_{\text{st}} \hat{Q}^x$, then we can take

$$\check{Q}^{n,b_n} \overset{D}{=} \hat{Q}^{D_{\tilde{\psi}_{b_n}^n}} \quad \text{or} \quad \check{Q}^{n,b_n} \overset{D}{=} \hat{Q}^{D_{\tilde{\psi}_1^n}},$$

respectively.

Lemma 4.6. *Under the above conditions,*

$$\lim_{\varepsilon \downarrow 0} \lim_{n \to \infty} P(\bar{T}_\infty^n \leq \varepsilon) = p_{\text{EXT}},$$

where p_{EXT} is as given in Corollary 3.1.

Proof. Let $\{b_n\}$ be any sequence of positive integers such that $b_n \to \infty$ and $b_n/n \to 0$ as $n \to \infty$. To prove the lemma, it suffices to show that $P(T_\infty^n \leq b_n) \to p_{\text{EXT}}$ as $n \to \infty$. In Section 3 it was shown that $T_\infty^n \leq_{\text{st}} Z_n$. Therefore,

$$\liminf_{n \to \infty} P(T_\infty^n \leq b_n) \geq \liminf_{n \to \infty} P(Z_n \leq b_n) \geq p_{\text{EXT}}. \tag{4.6}$$

For $n \geq 1$, let the branching process $\mathcal{B}_n^{b_n}$ be constructed as follows. Let $D_1, D_2, \ldots, D_n$ be independent and identically distributed according to D. Let there be a ancestors in the branching process $\mathcal{B}_n^{b_n}$. We assume that each initial ancestor has an independent lifetime distributed according to $\tilde{Q}$ and that each subsequent individual has an independent lifetime distributed according to $\check{Q}^{n,b_n}$. Furthermore, whilst alive, individuals reproduce at the points of independent Poisson point processes with rates $(1 - \delta_n)\lambda_n$, where $\delta_n = \sqrt{b_n/n}$. Thus, the offspring distributions of an initial ancestor and a subsequent individual are mixed Poisson; respectively, $\tilde{R}_n^{b_n} \sim \text{Po}((1 - \delta_n)\lambda_n \tilde{Q})$ and $\check{R}_n^{b_n} \sim \text{Po}((1 - \delta_n)\lambda_n \check{Q}^{n,b_n})$.

Let $\check{Z}^{n,b_n}$ denote the total progeny of the branching process $\mathcal{B}_n^{b_n}$. Then, since $(1 - \delta_n)\lambda_n \xrightarrow{P} E[D]$ and $\check{Q}^{n,b_n} \xrightarrow{D} V$ as $n \to \infty$, it follows that $\check{Z}^{n,b_n} \xrightarrow{D} Z$ as $n \to \infty$, where Z denotes the total progeny of the branching process $\mathcal{B}$ (see Section 3). Let p_{EXT}^{n,b_n} denote the extinction probability of the branching process $\mathcal{B}_n^{b_n}$. Then $p_{\text{EXT}}^{n,b_n} \to p_{\text{EXT}}$ as $n \to \infty$.

Let

$$A_n^{b_n} = \left\{ \frac{1}{n\lambda_n} \sum_{i \in \Psi_{b_n}^n} D_{\tilde{\psi}_i^n} \leq \delta_n \right\}.$$

From Corollary 4.1, $P(A_n^{b_n}) \to 1$ as $n \to \infty$. Therefore, whilst fewer than b_n of the initial susceptibles have been infected in the epidemic E_n, the epidemic E_n and the branching process

$\mathcal{B}_n^{b_n}$ can be coupled in such a way that, conditional upon $A_n^{b_n}$, every birth in the branching process has a corresponding successful infection in E_n. This is because

$$\check{Q}^{n,b_n} \leq_{st} Q^{D_{\check{\psi}_i^n}}$$

for $1 \leq i \leq b_n$ and the probability of an infective contact being made with a nonsusceptible is less than δ_n for $(1/n\lambda_n) \sum_{i \in \Psi_{b_n}^n} D_{\check{\psi}_i^n} \leq \delta_n$. Therefore,

$$\begin{aligned}
P(T_\infty^n \leq b_n) &\leq P(\{\check{Z}^{n,b_n} \leq b_n\} \cap A_n^{b_n}) + P((A_n^{b_n})^c) \\
&\leq P(\check{Z}^{n,b_n} \leq b_n) + P((A_n^{b_n})^c) \\
&\leq p_{EXT}^{n,b_n} + P((A_n^{b_n})^c).
\end{aligned} \quad (4.7)$$

The lemma follows immediately from (4.6) and (4.7) by taking the supremum limit in the latter equation.

Corollary 4.2. *For $\zeta = 0$,*

$$\lim_{n \to \infty} P(G_n) = 1 - p_{EXT}$$

and

$$[(\bar{T}_\infty^n, \bar{S}_\infty^n) \mid G_n] \xrightarrow{P} (\tau_0, \sigma_0) \quad \text{as } n \to \infty.$$

Proof. The proof is similar to [8, Theorem 3.12] and the details are thus omitted.

We can now focus on the limiting behaviour of a major epidemic. The main result is presented in Theorem 4.1 and follows standard arguments. The crux of the proof relies on the following lemma.

Lemma 4.7. *Suppose that $\bar{S}_\infty^n \xrightarrow{P} \sigma_\zeta$ as $n \to \infty$. Then, for any $\alpha, \beta \in \mathbb{R}$,*

$$|Z_n^{\alpha,\beta}(\bar{S}_\infty^n) - Z_n^{\alpha,\beta}(\sigma_\zeta)| \xrightarrow{P} 0 \quad \text{as } n \to \infty.$$

Proof. The proof is similar to [17, Lemma 5.4]. Fix $\varepsilon > 0$. By the theorem of total probability and Chebychev's inequality,

$$P(|Z_n^{\alpha,\beta}(\bar{S}_\infty^n) - Z_n^{\alpha,\beta}(\sigma_\zeta)| > \varepsilon) \leq \frac{1}{\varepsilon^2} E[var(Z_n^{\alpha,\beta}(\bar{S}_\infty^n) - Z_n^{\alpha,\beta}(\sigma_\zeta) \mid \bar{S}_\infty^n)]. \quad (4.8)$$

For any $s, t \geq 0$, note that the $(\chi_i(t) - \chi_i(s))(\alpha + \beta Q_i)$ are independent and identically distributed, whence

$$\begin{aligned}
var(Z_n^{\alpha,\beta}(t) - Z_n^{\alpha,\beta}(s)) &= var((\chi_1(t) - \chi_1(s))(\alpha + \beta Q_1)) \\
&\leq E[(\chi_1(t) - \chi_1(s))^2(\alpha + \beta Q_1)^2] \\
&= E[\chi_1(s \vee t)(\alpha + \beta Q_1)^2] - E[\chi_1(s \wedge t)(\alpha + \beta Q_1)^2]. \quad (4.9)
\end{aligned}$$

Therefore, from (4.8) and (4.9),

$$\begin{aligned}
P(|Z_n^{\alpha,\beta}(\bar{S}_\infty^n) - Z_n^{\alpha,\beta}(\sigma_\zeta)| > \varepsilon) &\leq \frac{1}{\varepsilon^2}\{E[\chi_1(\bar{S}_\infty^n \vee \sigma_\zeta)(\alpha + \beta Q_1)^2] \\
&\qquad - E[\chi_1(\bar{S}_\infty^n \wedge \sigma_\zeta)(\alpha + \beta Q_1)^2]\} \\
&= \frac{1}{\varepsilon^2}\{E[\{1 - \exp(-\{\bar{S}_\infty^n \vee \sigma_\zeta\}D_1)\}(\alpha + \beta Q_1)^2] \\
&\qquad - E[\{1 - \exp(-\{\bar{S}_\infty^n \wedge \sigma_\zeta\}D_1)\}(\alpha + \beta Q_1)^2]\}. \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (4.10)
\end{aligned}$$

Since $(\bar{S}^n_\infty \vee \sigma_\zeta, D_1, Q_1) \overset{P}{\to} (\sigma_\zeta, D_1, Q_1)$ as $n \to \infty$, from [11, Theorem 5.1 and Corollary 2],

$$\{1 - \exp(-\{\bar{S}^n_\infty \vee \sigma_\zeta\}D_1)\}(\alpha + \beta Q_1)^2 \overset{P}{\to} \{1 - \exp(-\sigma_\zeta D_1)\}(\alpha + \beta Q_1)^2.$$

Since $E[Q^2] < \infty$, from the bounded convergence theorem we find that

$$E[\{1 - \exp(-\{\bar{S}^n_\infty \vee \sigma_\zeta\}D_1)\}(\alpha + \beta Q_1)^2] \to E[\{1 - \exp(-\sigma_\zeta D_1)\}(\alpha + \beta Q_1)^2] \quad \text{as } n \to \infty.$$

Similarly,

$$E[\{1 - \exp(-\{\bar{S}^n_\infty \wedge \sigma_\zeta\}D_1)\}(\alpha + \beta Q_1)^2] \to E[\{1 - \exp(-\sigma_\zeta D_1)\}(\alpha + \beta Q_1)^2] \quad \text{as } n \to \infty.$$

Therefore, the right-hand side of (4.10) converges to 0 as $n \to \infty$, and the lemma is proved.

Theorem 4.1. *Suppose that there exists a $\zeta \geq 0$ such that $\sqrt{n}(a/n - \zeta) \overset{D}{\to} 0$ as $n \to \infty$. Then*

$$[\sqrt{n}(\bar{T}^n_\infty - \tau_\zeta) \mid G_n] \overset{D}{\to} Z^{1, r'(\sigma_\zeta)/(1 - a'(\sigma_\zeta))}(\sigma_\zeta) + \frac{r'(\sigma_\zeta)}{1 - a'(\sigma_\zeta)} I \quad \text{as } n \to \infty,$$

where $\sqrt{n}(\bar{S}^n_0 - \zeta E[\tilde{Q}]) = I_n \overset{D}{\to} I \sim N(0, \zeta \, var(\tilde{Q}))$ as $n \to \infty$.

Proof. The proof is similar to that of [17, Theorem 5.5], so only an outline is given here. By the mean value theorem, there exists a sequence $\{b^1_n\}$ such that $b^1_n \overset{P}{\to} \sigma_\zeta$ as $n \to \infty$ and such that

$$\sqrt{n}(\bar{T}^n_\infty - \tau) = Z^{1,0}_n(\bar{S}^n_\infty) + \sqrt{n} r'(b^1_n)(\bar{S}^n_\infty - \sigma_\zeta). \tag{4.11}$$

Simple algebraic manipulation of $\sqrt{n}(\bar{S}^n_\infty - \sigma_\zeta)$ gives

$$\sqrt{n}(\bar{S}^n_\infty - \sigma_\zeta) = I_n + Z^{0,1}_n(\bar{S}^n_\infty) + \sqrt{n} a'(b^2_n)(\bar{S}^n_\infty - \sigma_\zeta), \tag{4.12}$$

where $\{b^2_n\}$ is a sequence such that $b^2_n \overset{P}{\to} \sigma_\zeta$ as $n \to \infty$.

From Lemmas 4.7 and 4.2 and [11, Theorem 4.1], $Z^{0,1}_n(\bar{S}^n_\infty) \overset{P}{\to} Z^{0,1}(\sigma_\zeta)$. Therefore, since $a'(\sigma_\zeta) < 1$, by [11, Theorem 5.1, Corollary 2, and Theorem 4.4], we have

$$\sqrt{n} r'(b^1_n)(\bar{S}^n_\infty - \sigma_\zeta) = Z^{0, r'(\sigma_\zeta)/(1 - a'(\sigma_\zeta))}_n(\bar{S}^n_\infty) + \frac{r'(\sigma_\zeta)}{1 - a'(\sigma_\zeta)} I_n + B^1_n,$$

where $\{B^1_n\}$ is a sequence such that $B^1_n \overset{P}{\to} 0$ as $n \to \infty$. Finally, from (4.11) and Lemma 4.7,

$$\sqrt{n}(\bar{T}^n_\infty - \tau) = Z^{1, r'(\sigma_\zeta)/(1 - a'(\sigma_\zeta))}_n(\sigma_\zeta) + B^2_n + \frac{r'(\sigma_\zeta)}{1 - a'(\sigma_\zeta)} I_n + B^1_n, \tag{4.13}$$

where $\{B^2_n\}$ is a sequence such that $B^2_n \overset{P}{\to} 0$ as $n \to \infty$. The theorem follows from (4.13) and Lemma 4.2, since $Z^{1, r'(\sigma_\zeta)/(1 - a'(\sigma_\zeta))}_n(\sigma_\zeta)$ and I_n are independent.

5. Random-graph epidemic models

For an epidemic upon a mixed Bernoulli random graph, we shall bound the final size distribution of the random-graph epidemic by the final size distributions of a suitably defined VGSE and a perturbed VGSE. By showing that both these lower- and upper-bounding VGSEs have the same final size distributions, we can obtain the final size distribution of the epidemic upon a mixed Bernoulli random graph. We proceed by studying the perturbed VGSE. This is followed by a description of epidemics upon random graphs and an analysis of their final size distributions.

We shall consider a sequence of epidemics, $\{E_n^{\varepsilon_n}\}$, indexed by the total number of initial susceptibles, n, and a constant ε_n. For fixed $n \geq 1$ and $\varepsilon_n > -1$, we can construct the epidemics $E_n^{\varepsilon_n}$ and E_n on a common probability space $(\Omega, \mathcal{F}, \mathrm{P})$, with E_n constructed as in Section 2 and $E_n^{\varepsilon_n}$ constructed as follows. We assume that there a initial infectives and n initial susceptibles in $E_n^{\varepsilon_n}$, with the individuals labelled $-(a-1), -(a-2), \ldots, n$ as before. Let individual i in the epidemic $E_n^{\varepsilon_n}$ have life history (D_i, Q_i) for $1 \leq i \leq n$ and infective period $\tilde{Q}_i$ for $-(a-1) \leq i \leq 0$. Then, for $-(a-1) \leq i \leq n$ and $1 \leq j \leq n$, we assume that, during his infective period (of length Q_i or $\tilde{Q}_i$) in $E_n^{\varepsilon_n}$, individual i makes infective contact with individual j at the points of a homogeneous Poisson point process with rate $(1+\varepsilon_n)D_j/n$. Alternatively, we can assume that the infective period is of length $(1 + \varepsilon_n)Q_i$ and that the infection rate is D_j/n. Either way, the probability that individual i makes infective contact (at least once) with individual j is $1 - \exp(-(1+\varepsilon_n)Q_i D_j/n)$.

Let $T_\infty^{n,\varepsilon_n} = n\bar{T}_\infty^{n,\varepsilon_n}$ and $S_\infty^{n,\varepsilon_n} = n\bar{S}_\infty^{n,\varepsilon_n}$ respectively denote the final size and severity of the epidemic $E_n^{\varepsilon_n}$. Clearly, the epidemic E_n^0 is the epidemic E_n. We therefore have the following results if $\varepsilon_n \to 0$ as $n \to \infty$.

Theorem 5.1. *Suppose that $\varepsilon_n \to 0$ as $n \to \infty$. For all $n \geq 1$, if $a^n = a$ then*

$$T_\infty^{n,\varepsilon_n} \xrightarrow{\mathrm{D}} Z \quad as\ n \to \infty, \tag{5.1}$$

where Z is as defined in Section 3.

Suppose that $a/n \xrightarrow{\mathrm{P}} \zeta \geq 0$. Let σ_ζ and $\tau_\zeta = r(\sigma_\zeta)$ be as defined in Section 4. Then

$$\min\{|\bar{T}_\infty^{n,\varepsilon_n}|, |\bar{T}_\infty^{n,\varepsilon_n} - \tau_\zeta|\} \xrightarrow{\mathrm{P}} 0 \quad as\ n \to \infty. \tag{5.2}$$

Furthermore, if $\sqrt{n}\varepsilon_n \to 0$ and $G_n^{\varepsilon_n} = \{T_\infty^{n,\varepsilon_n} > \log n\}$, then

$$[\sqrt{n}(\bar{T}_\infty^{n,\varepsilon_n} - \tau_\zeta) \mid G_n^{\varepsilon_n}] \xrightarrow{\mathrm{D}} Z^{1,r'(\sigma_\zeta)/(1-a'(\sigma_\zeta))}(\sigma_\zeta) \quad as\ n \to \infty, \tag{5.3}$$

where $Z^{1,r'(\sigma_\zeta)/(1-a'(\sigma_\zeta))}(\sigma_\zeta)$ is as defined in Section 4.

Proof. The proof of this theorem requires only minor modifications of the proofs in Sections 3 and 4. Therefore, only an outline of the changes is given.

To prove (5.1), simply replace λ_n by $\lambda_n^{\varepsilon_n} = (1+\varepsilon_n)\lambda_n$, such that $\lambda_n^{\varepsilon_n} \xrightarrow{\mathrm{P}} \lambda$ as $n \to \infty$, in the proofs in Section 3.

The Sellke construction can then be used again, with $H_i \sim \mathrm{Exp}(D_i)$ and $(1+\varepsilon_n)Q_i$ respectively denoting the infective threshold and infective period of individual i. Let $R_n^{\varepsilon_n}(t)$ and $A_n^{\varepsilon_n}(t)$ (corresponding to $R_n(t)$ and $A_n(t)$) denote the total number of initial susceptibles infected and the corresponding sum of infective periods when each individual is exposed to t

units of global infective pressure in $E_n^{\varepsilon_n}$. Then $R_n^{\varepsilon_n}(t) = R_n(t)$ and $A_n^{\varepsilon_n}(t) = (1 + \varepsilon_n)A_n(t)$. Let $S_0^{n,\varepsilon_n} = (1 + \varepsilon_n)S_0^n$ denote the sum of the infective periods of the initial infectives. Then

$$\bar{S}_\infty^{n,\varepsilon_n} = \min\left\{s \geq 0 : s = (1 + \varepsilon_n)\left(\bar{S}_0^n + \frac{1}{n}A_n(s)\right)\right\}$$

(cf. (2.2)). Thus, it is trivial to adapt Lemma 4.3 to prove (5.2).

The adaptation of the lower-bounding branching process to the epidemic sequence $\{E_n^{\varepsilon_n}\}$ is straightforward by multiplying the birth rate by a factor of $1 + \varepsilon_n$. Therefore, Corollary 4.2 holds with $(\bar{S}_\infty^{n,\varepsilon_n}, \bar{T}_\infty^{n,\varepsilon_n}, G_n^{\varepsilon_n})$ in place of $(\bar{S}_\infty^n, \bar{T}_\infty^n, G_n)$. Thus, $[\bar{S}_\infty^{n,\varepsilon_n} \mid G_n^{\varepsilon_n}] \xrightarrow{\text{P}} \sigma_\zeta$ as $n \to \infty$ and Lemma 4.7 holds. Therefore, the proof of Theorem 4.1 can straightforwardly be used to prove (5.3) by replacing (4.12) with

$$\sqrt{n}(\bar{S}_\infty^{n,\varepsilon_n} - \sigma_\zeta) = \sqrt{n}(1 + \varepsilon_n)\left\{\bar{S}_0^n + \frac{1}{n}A_n(\bar{S}_\infty^{n,\varepsilon_n}) - \zeta\,\mathrm{E}[\tilde{Q}] - a(\sigma_\zeta)\right\}$$

$$= \sqrt{n}\varepsilon_n\left(\bar{S}_0^n + \frac{1}{n}A_n(\bar{S}_\infty^{n,\varepsilon_n})\right) + I_n + Z_n^{0,1}(\bar{S}_\infty^{n,\varepsilon_n}) + \sqrt{n}a'(b_n^2)(\bar{S}_\infty^{n,\varepsilon_n} - \sigma_\zeta)$$

$$(5.4)$$

and noting that the first term on the right-hand side of (5.4) converges in probability to 0 as $n \to \infty$.

We now turn our attention to random-graph epidemics. We shall consider a sequence of epidemics, $\{\hat{E}_n\}$, indexed by the total number of initial susceptibles, n, as $n \to \infty$. We assume that there are a initial infectives and n initial susceptibles. Let

$$(\tilde{C}_{-(a-1)}, \tilde{I}_{-(a-1)}), \qquad (\tilde{C}_{-(a-2)}, \tilde{I}_{-(a-2)}), \qquad \ldots, \qquad (\tilde{C}_0, \tilde{I}_0)$$

be independent and identically distributed according to $(\tilde{C}, \tilde{I})$. Let

$$(C_1, I_1), \qquad (C_2, I_2), \qquad \ldots, \qquad (C_n, I_n)$$

be independent and identically distributed according to (C, I). Let C and I respectively denote the connectivity and infectivity of an initial susceptible and let $\tilde{C}$ and $\tilde{I}$ respectively denote the connectivity and infectivity of an initial infective. For example, C could denote the sexual activeness of an individual in an epidemic of a sexually transmitted disease, or the connectedness of a computer in an epidemic of a network computer virus. There is assumed to be an underlying graph $\mathcal{G}_n$ on $n + a$ vertices which defines the acquaintance structure between the $n + a$ individuals in the population. The vertices are labelled $-(a - 1), -(a - 2), \ldots, n$ with vertex i corresponding to individual i. An edge exists between vertices i and j with probability $(1/n)C_iC_j \wedge 1$, where $C_k = \tilde{C}_k$, $-(a - 1) \leq k \leq 0$. In the epidemic $\hat{E}_n$, individuals i and j are said to be acquaintances if and only if there exists an edge between vertices i and j in $\mathcal{G}_n$. Individual i, whilst infective, makes infective contact with a given acquaintance j with probability I_i, independently of all other infective contacts. The epidemic and the random graph can be constructed in unison; see [16] and [17]. Specifically, whilst infective, individual i makes infective contact with a susceptible individual j with probability $((1/n)C_iC_j \wedge 1)I_i$. That is, the epidemic is constructed and then, if so desired, the random graph can be constructed retrospectively using the correct conditional distributions for the existence of edges between vertices. If the connectivity and infectivity random variables are

independent, then increased connectivity (activity) leads to increases in both the susceptibilities and infectivities of individuals. In particular, we can observe the effect of highly (sexually) active individuals such as prostitutes in the spread of a disease. Finally, let $\hat{T}_\infty^n = n\check{T}_\infty^n$ denote the final size of the epidemic $\hat{E}_n$.

Let $(D, Q) \overset{\mathrm{D}}{=} (C, CI)$. For $-(a-1) \leq i \leq 0$ set $\tilde{Q}_i = \tilde{C}_i \tilde{I}_i$, and for $1 \leq i \leq n$ set $(D_i, Q_i) = (C_i, C_i I_i)$. The epidemics E_n, $E_n^{\varepsilon_n}$, and $\hat{E}_n$ can be constructed upon a common probability space. For $-(a-1) \leq i \leq 0$, let the ith individuals in the epidemics E_n and $E_n^{\varepsilon_n}$ have infective periods of length $\tilde{Q}_i$ and let the ith individual in the epidemic $\hat{E}_n$ have connectivity and infectivity random variables $(\tilde{C}_i, \tilde{I}_i)$. For $1 \leq i \leq n$, let the ith individuals in the epidemics E_n and $E_n^{\varepsilon_n}$ have life histories (D_i, Q_i) and let the ith individual in the epidemic $\hat{E}_n$ have connectivity and infectivity random variables (C_i, I_i). For $-(a-1) \leq i \leq n$ and $1 \leq j \leq n$, let U_{ij} be independent and identically distributed according to $U \sim \mathrm{U}[0, 1]$. In E_n, $E_n^{\varepsilon_n}$, and $\hat{E}_n$, individual i, whilst infective, makes infective contact with individual j if and only if $U_{ij} \leq 1 - \exp(-Q_i D_j/n)$, $U_{ij} \leq 1 - \exp(-(1+\varepsilon_n)Q_i D_j/n)$, and $U_{ij} \leq ((1/n)C_i C_j \wedge 1)I_i$, respectively.

Suppose, for $0 < \varepsilon_n < 1$ and all i, $-(a-1) \leq i \leq n$, and j, $1 \leq j \leq n$, that $(1/n)C_i C_j \leq \varepsilon_n$. Then, since $0 \leq I_i \leq 1$, $Q_i = C_i I_i$, and $D_j = C_j$, it follows that

$$1 - \exp\left(-\frac{Q_i D_j}{n}\right) \leq \left(\frac{1}{n}C_i C_j \wedge 1\right)I_i \leq 1 - \exp\left(-(1+\varepsilon_n)\frac{Q_i D_j}{n}\right). \tag{5.5}$$

Therefore, using the above construction of the epidemics, we have

$$T_\infty^n \leq \hat{T}_\infty^n \leq T_\infty^{n, \varepsilon_n}. \tag{5.6}$$

Thus, the final size of the epidemic $\hat{E}_n$ is sandwiched between the final sizes of the epidemics E_n and $E_n^{\varepsilon_n}$. Therefore, although the epidemics E_n and $\hat{E}_n$ are very different in appearance, we can show that they have the same limiting asymptotic final size distributions.

Theorem 5.2. *Suppose that there exists a $\delta > 0$ such that $\mathrm{E}[C^{2+\delta}]$, $\mathrm{E}[\tilde{C}^{2+\delta}] < \infty$. Then*

$$\hat{T}_\infty^n \overset{\mathrm{D}}{\to} Z \quad as\ n \to \infty, \tag{5.7}$$

where Z is as defined in Section 3.

Suppose that $a/n \to \zeta \geq 0$. Let τ_ζ and σ_ζ be as defined in Section 4. Then $\check{T}_\infty^n \overset{\mathrm{P}}{\to} \tau_\zeta$ as $n \to \infty$, for $\zeta > 0$, and

$$\min\{|\check{T}_\infty^n|, |\check{T}_\infty^n - \tau_\zeta|\} \overset{\mathrm{P}}{\to} 0 \quad as\ n \to \infty, \tag{5.8}$$

for $\zeta = 0$. Let $\hat{G}_n = \{\hat{T}_\infty^n > \log n\}$ denote the event that a global epidemic occurs in $\hat{E}_n$. Then $\mathrm{P}(\hat{G}_n) \to 1$ for $\zeta > 0$ and, for $\zeta = 0$, $\mathrm{P}(\hat{G}_n) \to 1 - p_{\mathrm{EXT}}$ as $n \to \infty$, where p_{EXT} is as given in Corollary 3.1.

Furthermore, if there exists a $\delta > 2$ such that $\mathrm{E}[C^{2+\delta}]$, $\mathrm{E}[\tilde{C}^{2+\delta}] < \infty$, then

$$[\sqrt{n}(\check{T}_\infty^n - \tau_\zeta) \mid \hat{G}_n] \overset{\mathrm{D}}{\to} Z^{1, r'(\sigma_\zeta)/(1-a'(\sigma_\zeta))}(\sigma_\zeta) + \frac{r'(\sigma_\zeta)}{1 - a'(\sigma_\zeta)}I \quad as\ n \to \infty, \tag{5.9}$$

where $Z^{1, r'(\sigma_\zeta)/(1-a'(\sigma_\zeta))}(\sigma_\zeta)$ and I are as defined in Section 4.

Proof. From the results of Sections 3 and 4, Theorem 5.1, and (5.6), it suffices to show that there exists a suitable sequence $\{\varepsilon_n\}$ such that, with probability tending to 1 as $n \to \infty$, (5.5) holds for all i, $-(a-1) \leq i \leq n$ and j, $1 \leq j \leq n$.

For $\kappa > 0$, let $A_n^\kappa = \bigcap_{i=-(a-1)}^n \{C_i < n^\kappa\}$, where $C_i = \tilde{C}_i$, $-(a-1) \leq i \leq 0$. By Markov's inequality, $\mathrm{P}((A_n^\kappa)^c) \to 0$ as $n \to \infty$, for $\kappa > 1/(2+\delta)$. Therefore, let $1/(2+\delta) < \kappa < \frac{1}{2}$ and $\varepsilon_n = n^{2\kappa-1}$. Then (5.7) and (5.8) follow immediately. Equation (5.9) is proved similarly, provided that $\delta > 2$, by letting $1/(2+\delta) < \kappa < \frac{1}{4}$ and $\varepsilon_n = n^{2\kappa-1}$.

Finally, note that, for the Bernoulli random-graph epidemic, $\hat{Q}^x = xI$ and, so, $\hat{Q}^x \leq_{\mathrm{st}} \hat{Q}^y$ for all $x < y$. Therefore, the assumption in the prelude to Lemma 4.6 holds.

Alternative models can be used to fix the probability of an edge between vertices i and j in the mixed Bernoulli random graph. In particular, in [12, Section 3] the probability that an edge exists between vertices i and j is $C_i C_j/(n + C_i C_j)$. Theorem 5.2 holds since, with probability tending to 1 as $n \to \infty$, we can bound the resulting random-graph epidemic using $E_n^{-\varepsilon_n}$ and $E_n^{\varepsilon_n}$.

6. Discussion

There are a number of obvious extensions of both the VGSE model and the random-graph epidemic model. We briefly discuss two of these below.

Firstly, it would be interesting to consider a multitype VGSE model; see, for example, [5] and [17]. This is particularly relevant for the vaccination models of [10], where an individual's (susceptible and infective) response to vaccination is variable. In such a case we could consider a two-type model with the types corresponding to vaccinated and unvaccinated individuals. Extending the results of Sections 3 and 4 to include more than one type of individual should be relatively straightforward following [5] and [17].

Secondly, coupling an epidemic upon a random graph to a VGSE can prove fruitful in examining epidemics upon graphs with small cliques. These cliques could correspond to small, completely mixing social groups such as households [1], [2]. No previous analysis of epidemics upon (random) graphs can incorporate the effect of cliques. However, the GSE household epidemic model is well understood (see, for example, [9] and [2]), and extending the current results to a household epidemic model with variable global (and possibly local) susceptibility should not present any real difficulties.

Acknowledgements

References

[1] ANDERSSON, H. (1998). Limit theorems for a random-graph epidemic model. *Ann. Appl. Prob.* **8,** 1331–1349.

[2] ANDERSSON, H. (1999). Epidemic models and social networks. *Math. Scientist* **24,** 128–147.

[3] BALL, F. (1985). Deterministic and stochastic epidemics with several kinds of susceptibles. *Adv. Appl. Prob.* **17,** 1–22.

[4] BALL, F. AND BRITTON, T. (2005). An epidemic model with exposure-dependent severities. *J. Appl. Prob.* **42,** 932–949.

[5] BALL, F. AND CLANCY, D. (1993). The final size and severity of a generalised stochastic multitype epidemic model. *Adv. Appl. Prob.* **25,** 721–736.

[6] BALL, F. AND CLANCY, D. (1995). The final outcome of an epidemic model with several different types of infective in a large population. *J. Appl. Prob.* **32,** 579–590.

[7] BALL, F. AND LYNE, O. (2001). Stochastic multitype SIR epidemics among a population partitioned into households. *Adv. Appl. Prob.* **33,** 99–123.

[8] BALL, F. AND NEAL, P. (2003). The great circle epidemic model. *Stoch. Process. Appl.* **107,** 233–268.

[9] BALL, F., MOLLISON, D. AND SCALIA-TOMBA, G. (1997). Epidemics with two levels of mixing. *Ann. Appl. Prob.* **7,** 46–89.

[10] BECKER, N. AND STARCZAK, D. (1998). The effect of random vaccine response on the vaccination coverage required to prevent epidemics. *Math. Biosci.* **154,** 117–135.

[11] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. John Wiley, New York.

[12] BRITTON, T., DEIJFEN, M. AND MARTIN-LÖF, A. (2006). Generating simple random graphs with prescribed degree distribution. *J. Statist. Phys.* **124,** 1377–1397.

[13] HOLST, L. (2001). Extreme value distributions for random coupon collector and birthday problems. *Extremes* **4,** 129–145.

[14] LEFÈVRE, C. AND PICARD, P. (2005). Nonstationarity and randomization in the Reed–Frost epidemic model. *J. Appl. Prob.* **42,** 950–963.

[15] LEFÈVRE, C. AND UTEV, S. (1999). Branching approximation for the collective epidemic model. *Methodology Comput. Appl. Prob.* **1,** 211–228.

[16] NEAL, P. (2003). SIR epidemics on a Bernoulli random graph. *J. Appl. Prob.* **40,** 779–782.

[17] NEAL, P. (2006). Multitype randomised Reed–Frost epidemics and epidemics upon random graphs. *Ann. Appl. Prob.* **16,** 1166–1189.

[18] SCALIA-TOMBA, G. (1985). Asymptotic final size distribution for some chain-binomial processes. *Adv. Appl. Prob.* **17,** 477–495.

[19] SCALIA-TOMBA, G. (1990). On the asymptotic final size distribution of epidemics in heterogeneous populations. In *Stochastic Processes in Epidemic Theory* (Lecture Notes Biomath. **86**), Springer, New York, pp. 189–196.

[20] SELLKE, T. (1983). On the asymptotic distribution of the size of a stochastic epidemic. *J. Appl. Prob.* **20,** 390–394.

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